

Language and Morality: Evolution, Altruism, and Linguistic Moral Mechanisms

Joseph W. Poulshock

**A thesis submitted in fulfillment of requirements for the degree of
Doctor of Philosophy**

**to
Theoretical and Applied Linguistics
School of Philosophy, Psychology and Language Sciences
University of Edinburgh**

February 2006

© Copyright 2006
By
Joseph W. Poulshock

ABSTRACT

This thesis inquires into how human language relates to morality -- and shows the ways language enables, extends, and maintains human value systems. Though we ultimately need to view the relation between language and morality from many different perspectives -- biological, psychological, sociological, and philosophical - the approach here is primarily a linguistic one informed by evolutionary theory.

At first, this study shows how natural selection relates to the problem of altruism and how language serves human moral ontogeny. Subsequently, the argument demonstrates how language helps enable cultural group selection. Moreover, as language helps influence human behavior in an altruistic direction beyond in-group non-kin (helping facilitate cultural group selection), we also consider how language can help facilitate altruistic behavior towards out-group non-kin. This therefore raises the prospect of a limited moral realism in a world of evolutionary processes.

With these issues and possibilities in mind, we consider and analyze the properties of language that help extend human morality. Specifically, discussion covers how recursion, linguistic creativity, naming ability, displacement, stimulus freedom, compositionality, cultural transmission, and categorization extend moral systems. Moreover, because language so broadly influences morality, the inquiry extends into how linguistic differences (specifically between English and Japanese) might also cause subtle differences in moral perception between Japanese and English speakers.

Lastly, we consider how moral ideas might take on a life of their own, catalytically propagating in degrees dependent and independent of human intention. That is, we consider how ideas might become memetic. After considering the serious problems

of memetics, this approach employs a linguistic version of memetic theory and considers how psychological, social, and linguistic constraints may cause moral memes to attain a memetic state and spread by an independent or semi-independent replicator dynamic. Thus, some moral ideas that we possess through language may actually possess us.

Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Joseph W. Poulshock

ACKNOWLEDGEMENTS

Here is a brief list of acknowledgements where I gratefully confess to a long period of deep indebtedness to colleagues, family, and friends. Heartfelt thanks go to:

- Jim and Sue Hurford, Jim for excellent supervision, and Sue for the most pleasant coffee, tea, cookies, and camaraderie.
- Simon Kirby, for challenging me, believing in me, and for the music.
- Miriam Meyerhoff, for positive, insightful, and wise supervision.
- To the LEC Gang: to Andrew Smith for the loudest laugh on the planet, and for invaluable advice; to Kenny Smith, for teaching me Scottish and for answering *all* my questions; to Jimmy Yamauchi, for helpful support and profound insight into single malt; to Henry Brighton, for increasing my artificial intelligence and keeping me on my toes; to Jelle Zuidema, for being someone to look up to and think deeply with; to Anna Parker, Paul Vogt, Viktor Tron, and the many others who helped make the LEC such a cool place to work.
- To The People of the Common Room: Cassie Mayo, Rob Clark, Dan Wedgwood, Michael Bennett, Mits Ota, Lynda Hutchinson, Francesca Filiaci, Maria Luisa Flecha-Garcia, and many others who made eating, drinking, talking, laughing, and thinking in Edinburgh so enjoyable.
- To colleagues at TCU: to Stephen Franklin -- the ideal mentor; Noyuri Sugitani, for translating the surveys and for constant encouragement; to Hisakazu Inagaki, Nobuyuki Takahashi, and KCI, for interest in this work; to Akio Ito and Takanori Kobayashi for institutional support. I am also grateful to TCU for 2 sabbaticals, and for giving me the time to complete this task.
- To Hanako Misao, at Saint Luke's International Hospital, for help with statistics.
- To Richard Hamilton, for reminding me of the value this research; to Jeff Schloss for inspiring visits at Westmont; to Jay Richards and Steve Meyer for exemplifying gentlemanliness and scholarship; to the Templeton Foundation for indirectly supporting this research.
- To BB and the support of CJOIS during the final stages of this writing; to Scott Maddan, for hospitality and friendship that supported this work.
- To my family: thank you all literally -- for everything.

CONTENTS

Abstract	v
Declaration	vii
Acknowledgements	ix
List of Figures	xv
Chapter 1	1
Introduction	1
1.1 Language, morality, evolution, and social groups	1
1.2 The Problem of Altruism Language and Social Groups	3
1.3 Linguistically Encoded Morality and Cultural Evolution	4
1.4 Summary of this Thesis	5
Chapter 2	7
On the Linguistic Basis of Morality	7
2.1 Language an Evolutionary Prerequisite to Morality	7
2.2 The Problem of Human Morality Explained by Language	9
2.2.1 The Assessment Function	9
2.2.2 The Enforcement Function	10
2.2.3 The Shared Representation Function	11
2.2.4 The Cooperation Function	12

2.2.5 The Group Fitness and Conflict Function	14
2.3 The Degrees to which Human Morality is Possible Without Language.....	16
2.4 The Parallel Between Linguistic and Moral Development in Children.....	21
2.5 Moral Ontogeny: Broader Considerations	25
2.6 Language and Morality for Feral or Abused Children.....	30
2.7 Language Innateness and Moral Ontogeny	33
2.8 Apparent Altruistic or Moral Behavior in Non-Human Species.....	35
2.9 Animals, Language, Altruism, and Morality.....	38
2.10 Symbolic Language, Empathy, Reciprocity and Moral Codes.....	41
2.11 Cooperation, Fairness, Trust, Play, and Language	47
2.12 Natural Evil and its Relation to Morality and Altruism	50
Chapter 3	56
Language, Morality, and Group Selection	56
3.1 Altruism and Morality: A Central Problem for Evolutionary Theory	56
3.2 Solutions for the Problem of Altruism?	57
3.3 Explaining Real Altruism as Anomalous.....	60
3.4 Justifying an Appeal to Group Selection	64
3.5 The Relation Between Group Selection and Language	70
3.5.1 How does language influence groupishness?	70
3.6 How does language influence morality?.....	72
3.7 Language and the Transmission of Moral Codes	76
3.8 Group Identity and Code Regulation of Group Behavior	78
3.9 The Dark Side of Group Selection.....	80
3.10 Introducing Virtual Relatedness	82
3.11 The Reciprocity Principle.....	87

3.12 Group Selection, Morality and Religion	89
3.13 Conclusions	92
Chapter 4.....	95
Language and Morality Beyond Group Selection.....	95
4.1 Real Altruism Directed to Out-group Members.....	95
4.2 Group Selection is Selfish Selection Between Groups.....	97
4.3 Atypical Altruism Transcending GS and SGT	99
4.4 TFT Reciprocity Between Groups and Radical Altruism.....	101
4.5 Radical Altruism Beyond TFT Between Group Reciprocity.....	106
4.6 Getting at the Root of Maladaptive Morality: Nonrealism.....	108
4.7 The Relation Between Moral Realism and Human Language.....	112
4.8 The Role of Language in Developing Altruistic Personality	117
4.9 The Altruistic Perspective and Linguistic Ability.....	124
Chapter 5.....	130
The Properties of Language that Extend Morality	130
5.1 Linguistic Properties of Language that Enable or Enhance Morality	130
5.2 Similarities and Differences Between UG and Universal Morality?	130
5.2.1 Empirical Support for UG.....	131
5.3 Evidence for UG Correlated to the Human Moral Capacity	134
5.4 Recursive Linguistic Creativity Enhances Morality	138
5.5 Creativity, Naming, and Morality	141
5.6 Linguistic Displacement Enhances Morality	144
5.7 The Influence of Stimulus Freedom on Morality.....	146
5.8 Cultural Transmission and Moral Values	148

5.9 Language, Categorization, and Morality.....	149
5.10 Design Features of Language, Identity, and Morality.....	154
5.11 Design Features of Language and Virtual Relatedness.....	156
Chapter 6	161
The Influence of Language on Moral Perception.....	161
6.1 Linguistic Relativity Light in the Realm of Values and Morals	161
6.1.1 Justifying Research into Linguistic Relativity and Moral Perception....	163
6.1.2 Differences Between English and Japanese	164
6.1.3 Semantic Differences Between English and Japanese	165
6.1.4 The Intuitions of Bilingual Speakers.....	166
6.1.5 Empirical Evidences for Linguistic Relativity.....	168
6.1.6 Linking Linguistic Relativity and Moral Perception	170
6.2 Methodology of the Moral Perception/Dilemma Experiment.....	170
6.2.1 Preparation and Piloting	171
6.2.2 Results of the Preliminary/Pilot Survey	172
6.2.3 Methodology and Results of the Secondary Survey	173
6.2.4 Discussion and Results of the Moral Dilemma Survey.....	173
6.2.5 Further Statistical Interpretation of Questions 3 and 12 (Secondary Survey).....	176
6.3 Additional Reflections on Linguistic Relativity and Morality	177
6.4 Three Summations	181
6.4.1 The Fitness of Concepts	182
6.4.2 Translation and Underdetermined Semantic Spaces.....	182
6.4.3 Linguistic Relativity and Ethnography.....	184
6.5 Conclusion.....	186
Chapter 7	189
Language, Evolution, Morality, and Memes.....	189
7.1 Language, Evolution, Morality and Cultural Transmission Theory.....	189
7.2 Basic Background: Where are the Real Memes?	193
7.2.1 Basic Background: Memetic Diplomacy, Pragmatism, and Mechanisms	194
7.2.2 Background of Radical Memetics.....	196
7.3 Definitional Problems with Radical Memetics.....	197

7.4 Misguided and Misinformed Aspects of Memetics	200
7.4.1 Scenario: If Memes Had a Mind of their Own.....	203
7.5 Beyond Memetic Madness.....	203
7.6 The Potential of Memetics in Metaphysics and Science	205
7.6.1 Neuromemetics.....	206
7.6.2 The Flaws of Neuromemetics.....	208
7.6.3 Linguistic Problems with Neuromemetics	211
7.7 Lingua-memetics: Fitness-enhancing information in a memetic state	214
7.7.1 Looking for Memes in Language	215
7.8 Potential Examples of Linguistically Based Moral Memes.....	218
Chapter 8.....	227
Conclusions	227
8.1 Summary	227
8.2 Future Research.....	230
Appendix A & B.....	232
Bilingual Moral Dilemma Surveys.....	232
References.....	238

LIST OF FIGURES

Figure 3.1: Space of Biologically Related and Unrelated individuals.....	84
Figure 4.1: Coverage of Theories of Altruism	96
Figure 4.2: Diagram of Support for Maintaining Radical Altruism	103
Figure 5.1: The Golden Rule and Semantic Primitives	158
Figure 6.1: You/I Pronouns in Japanese.	165
Figure 6.2: J/E Bilinguals Intuitions Related to Linguistic Relativity.....	166

CHAPTER 1

INTRODUCTION

1.1 Language, morality, evolution, and social groups

Evolutionary linguistics forms a relatively new field of research that approaches the subfields of linguistics from a Darwinian perspective. From this viewpoint, linguists can focus on how natural selection, cultural evolution, and other causal pressures, mechanisms, and processes may have influenced the origin, evolution, and survival of various linguistic behaviors and features (Hurford et al., 1998; Knight et al., 2000; Wray, 2002). This field of inquiry is diverse -- requiring a multidisciplinary approach to investigate exceptionally challenging questions, such as the origin of language, which scholars consider a major transition in evolution that requires complex and painstaking explanation (Maynard-Smith and Szathmary, 1999; Wray, 2002).

With these fundamental questions stands the related evolutionary problem of altruism and how human language interacts with the origin, evolution, and maintenance of altruistic and moral behavior in human groups. Does morality require language? Could humans have developed a morality without language? Could we have developed a language without morality? What is the evolutionary relationship between language, altruism, and morality? These language-related questions connect to the general Darwinian problem of altruism (Hamilton, 1964; Trivers, 1971; Dawkins, 1976; Sober and Wilson, 1998) in ways that may transform and invigorate discussion on this topic. The orthodox evolutionary view claims that organisms best adapted to their environments experience the highest

levels of reproductive fitness and evolve through natural selection. This process is entirely selfish as it only concerns individual organisms that reproduce the genotypes that in turn produce the most well adapted phenotypes (Dawkins, 1976; 1982). This makes altruism evolutionarily disadvantageous if it does not benefit the altruist's genes.

Thus, in the mainstream view, genetically selfish organisms focus on self-reproduction and do not express real altruistic behavior towards non-relatives. Consistent with this logic, one accepted view explains that genuine altruistic behavior in humans and other species exists only in appearance; that is, altruistic behavior directed toward kin members (Hamilton, 1964) is actually selfish because it promotes the reproductive fitness of the altruist's genes that exist in the kin member. A second form of apparent altruism directed at non-kin is called reciprocal -- or tit-for-tat -- altruism (Trivers, 1971). This claims that altruistic behavior directed toward non-kin only occurs when the recipient of the altruist's efforts reciprocates, and this makes reciprocal altruism somewhat selfish. Moreover, if the receiver does not reciprocate, then the altruist will cease to be altruistic to that "defector."

Since reciprocal altruism expects a return on its efforts, it still benefits the altruist's genes and therefore does not contradict selfish gene theory. Nevertheless, some researchers claim (de Waal, 1996; Sober and Wilson, 1998; Sober and Wilson, 2000; Bekoff, 2002) that a genuine, as opposed to apparent altruism, exists that transcends genetic selfishness and therefore selfish gene explanations. We may describe this kind of altruism minimally as behavior whose cost is greater than the reciprocal benefit the individual altruist may receive in return for the altruistic act, and we may call this behavior disinterested altruism, self-sacrificial altruism, high-cost altruism, real altruism, or genuine altruism. These scholars claim that a truly unselfish and costly form of altruism exists in the natural world: altruism that fails to benefit the altruist's genes and apparently transcends the limits of selfish gene theory.

At this point, we need to clarify terms. First, any kind of selfish altruism or real altruism may express itself in ways that we might call unethical. That is, it is

possible to behave altruistically in unethical ways; for example, a student could risk his standing at school and steal the answers to a test for his group of friends. A soldier, at risk of death, can inadvertently or purposely take the lives of non-combatant civilians for his tribe's economic or territorial benefit; a terrorist can sacrifice his life and the lives of innocent women and children for his political cause. Thus, in this thesis, we may use the terms morality and altruism somewhat interchangeably in a purely descriptive sense. However, to differentiate them, the term "selfish altruism" will refer to various forms of other regarding behavior that include a fitness cost with a equal or greater return benefit to the altruist, as in kin selective altruism or reciprocal altruism.

"Genuine altruism," "real altruism," and similar terms will refer to non-kin-directed, non-reciprocal altruism, which would naturally incur a fitness cost but with less or no return benefit to the altruist. The term "morality" will descriptively refer to how individuals or groups define or categorize particular altruistic or non-altruistic behaviors as good or bad, and these evaluations are connected to the concepts of "must" and "ought." That is, we *must* not do that bad behavior, and we *ought* to do that good behavior. Hence, Group A might label Behavior X as altruistic and good, but Group B might label the same behavior as bad. For the descriptive purposes of this thesis, the categories of good and bad emerge from the subjects in question and not necessarily from the perspective of the researcher or the reader. Lastly, we may find it necessary to consider how this research relates to our own prescriptive norms, and for this perspective, the term "ethics" will be used.

1.2 The Problem of Altruism Language and Social Groups

Generally, those who accept that genuine altruism actually exists in nature will often appeal to group selection or multi-level selection pressures. Many scholars, however, deem this problematic because until recently group selection was considered utterly discredited as an explanatory tool (Hamilton, 1964; Williams, 1966; Dawkins, 1976; Trivers, 1985). Nonetheless, mostly thanks to the work of Sober and Wilson (1998), group selection (hereafter GS), though still controversial, has experienced renewed rigor, broader consideration, and productive empirical testing (Dicks, 2000; Sober and Wilson, 2000; Bekoff, 2002).

One problem with GS concerns how a group of genetically selfish individuals can coalesce into a unit to the degree that it would undergo group selection pressures. In other words, how can a group of individuals who naturally act in their own genetic self-interest begin to behave altruistically to non-kin members thus incurring a cost to their own reproductive fitness? And how would this process allow the group in question to coalesce to the degree that it would begin to undergo a GS process? With this in mind, this thesis claims and explains how human language provides a key for human groups to promote genuine altruism, enabling them to unify and therefore outcompete adversaries, harsh environmental conditions, and destructive concepts that would promote selfishness over genuine altruism and damage the group's fitness.

1.3 Linguistically Encoded Morality and Cultural Evolution

Looking at language as a key for understanding the origin, evolution, and maintenance of altruism and morality assumes the most basic question of how language originated in the first place. Nevertheless, this thesis still makes the bold claim that once complex language exists, human groups can avail themselves to a new kind of cultural evolution -- that through the aid and support of language -- humans can overcome putative genetic self-interest, produce genuine altruism, and avail themselves to group selection pressures. Interestingly, this process implies that genuine altruistic behavior that benefits group fitness also benefits individuals, but in a very indirect way. Based on a linguistically encoded abstract value system, Person X behaves altruistically to various non-kin, and generally does not receive a direct tit-for-tat benefit. However, since the linguistically encoded abstract value system may promote real altruism in a sufficient number of group members in an adequate number of situations, an indirect benefit may return to individuals in the protection and support that the group provides. Hence, individual benefit may come from this kind of genuine altruism.

However, this does not appear to be selfish gene altruism because a genuine altruism would benefit the group (not just the individual), helping it to coalesce and undergo multi-level selection pressures. In addition, the group does not need to be genetically defined; in fact, social groups can be genetically very heterogeneous.

Moreover, the risks involved could potentially lead to ultimate sacrifice that would give no indirect return benefit to at least some altruists. Incidentally, we should note that group oriented altruism may take many forms, including benevolent behavior towards non-kin, but also a ruthless kind of altruism that promotes the fitness of the in-group and the non-fitness of the out-group. These group conflict issues (the dark side of group selection) will be discussed at a later point.

Nevertheless, this thesis contends that human language helps enable value systems, which we may sometimes consider unethical, but that still facilitate GS and group fitness over and against selection pressures that would be detrimental to the fitness of groups.

1.4 Summary of this Thesis

In sum, with this general background in mind, this thesis attempts to answer the following overarching questions -- some of which are explained above, and others that will derive from or relate to the logical outworking of these ideas.

1. Does morality require language?
2. What role does language play in human altruistic and moral ontogeny?
3. Can language-based norms facilitate group selection?
4. Can moral-linguistic systems enable altruistic behavior that transcends group selection?
5. What design features of language enable and extend human altruism and morality?
6. Can language influence moral attitudes and perception?
7. Can some moral concepts possess meme-like traits that cause them to replicate between human hosts in degrees independent of human intention?

Darwin, in Chapter Four of *The Origin of the Species*, maintained that the human moral sense exists as the most important difference between humanity and the lower animals. Moreover, leading evolutionary linguists concur that language stands out boldly as a uniquely human characteristic (Pinker, 1994; Deacon, 1997; Hurford et al., 1998; Calvin and Bickerton, 2000; Kirby, 2000; Jackendoff, 2002). Hence, the uniqueness and therefore significance of both human morality and language mingle two vital, intriguing, and interconnected traits that obviously merit intensive study.

CHAPTER 2

ON THE LINGUISTIC BASIS OF MORALITY

2.1 Language an Evolutionary Prerequisite to Morality

The main premise of this thesis proposes that morality -- human morality -- requires language. Related to this, Sober and Wilson (Sober and Wilson, 2000: 186, 205) claim that due to the incompleteness of current theories, "evolutionary altruism [poses] a fundamental problem for the theory of natural selection." Therefore, they advocate group selection (GS) to help explain the contours of altruism and morality that mainstream solutions do not fully explain. The second basic premise of this study suggests that human language provides a means for GS to operate on human groups. In short, this thesis makes two main claims that *morality requires language, and linguistically enabled morality helps facilitate GS.*

A third premise concerns the existence of real altruism. Along these lines, Richard Dawkins (1976), the leading selfish geneticist who continues to oppose GS (cp. Dicks, 2000) asserts that real altruism has no place in human biological nature (1976: 3, 150, 215). However, Dawkins also proposes that memes (defined as the cultural analog to genes) help create a moral code in culture that enables humans to rebel against the tyranny of our selfish replicators, thus entailing that we can produce real, cooperative, even high cost altruism. That is, something produced by human nature, namely factors in language and culture enable us to transcend our biological nature. The serious problems with memetics will be critiqued in Chapter 7. However, linguistically based moral codes often oblige people towards genuine altruism, and even if we only rarely meet such obligations, this implies that where

these codes exist, selfish gene explanations of altruism (reciprocal altruism and kin selection) may no longer stand as the *only* theories that explain altruism and morality.

Therefore, if moral codes that explicitly promote real altruism exist in culture, and even if these codes only rarely succeed, this counters any logical restriction against GS and genuine altruism in human groups. That is, the cultural presence of linguistically based moral codes that even negligibly enable us to transcend the power of selfish genes undercuts preemptions that define group selection and genuine altruism out of the explanatory picture. The point here is that the gene's eye view can delimit and bias our view such that we cannot see real altruism if it exists.

Hence, from the selfish-gene view, we can mistakenly label potentially genuine altruism as selfishness or apparent altruism. Nevertheless, culture and language are inexorably connected, and culture equipped with the power of language appears to allow for a Lamarckian kind of evolution of inherited characteristics (Gould, 2001). If cultural evolution is Lamarckian, then linguistically based norms encoded in culture could also enable human altruism to transcend more basic Darwinian selfish genic processes -- as Dawkins himself suggests. And this could open the door to additional selection pressures, including GS. Dawkins would deny that his proposal implies GS, but GS potentially emerges from his thinking because he claims that memes allow for real altruism.

We can even discuss ways of deliberately cultivating and nurturing a pure, disinterested altruism -- something that has no place in nature, something that has never existed before in the whole history of the world. We are built as gene machines and cultured as meme machines, but we have the power to turn against our creators. We, alone on earth, can rebel against the tyranny of the selfish replicators. (Dawkins, 1976: 215)

The serious problems of memes aside, humans create moral codes that advocate real altruism; if these codes can even partly promote real altruism between non-kin within a group, they can also promote group selection. This stands as one of my basic claims, which leads to questions about how we can thoroughly develop the logic for this proposal, and most importantly, how we can find empirical evidence for it.

2.2 The Problem of Human Morality Explained by Language

The evolutionary problem of altruism is this: when an organism behaves altruistically to a conspecific, it risks reproductive success. Thus, truly unselfish organisms can altruistically behave themselves right out of the gene pool, making such generous behavior evolutionarily unstable, especially if we do not consider groups as part of the explanatory picture where the group would benefit from the sacrificial behavior. Moreover, since biological predispositions cause organisms to behave in genetically selfish ways, a leveraging mechanism might increase the probability of altruistic behavior, and language is one such potential lever. Many overlapping reasons exist for how a lever of language might promote altruism, and the following section describes five such reasons in terms of functions that are best (or only) solved by language.

2.2.1 The Assessment Function.

It is hard to assess morality in species without language (Bernstein, 2000), but language helps us access and assess moral thinking. Perhaps we can only assess norms, beliefs, and values through language and questionnaires, leaving other species outside the realm of such inquiry. Thus, assessment refers to checking for the existence of particular moral values and sometimes comparing behavior to language-based standards. To be sure, researchers can observe and analyze substrate morality in non-human species (de Waal, 1996; Bekoff, 2002), but the lack of language in nonhuman species limits the analysis of their moral behavior. We cannot ask non-linguistic species why they behave in certain ways, or whether they think a particular behavior is selfish, unselfish, good, or bad, or whether they think they *must* behave this way, or that they *ought* not behave that way.

This might not be the case regarding language-trained apes, but these animals are no longer non-linguistic because of their training, and if language enables us to find out more about the substrate morality of these species, this simply supports the point under consideration. Nevertheless, lack of language brings a benefit by allowing researchers to focus on fewer variables, as well as consider what

difference the absence or presence of language makes on moral behavior. On the other hand, through linguistic means, such as a survey, for example, it becomes much easier to access and assess thinking and behavior related to morality.

Moreover, language and gossip permit humans to articulate norms and monitor altruists and non-altruists who may be present or absent from the observers (Boehm, 2000). Hence, language not only enables us to articulate norms in a cost-efficient manner -- with a few words, such as "Thou shalt not..." or "Do as you would be done to," but when group members articulate and agree upon norms, they can compare actions with the standards set (Sober and Wilson, 2000).

Regarding assessment, since humans use language to depict, codify, and create moral values, morality strongly depends on language, and this contingency between language and morality makes assessing *human morality* dependent on language. If altruism-promoting codes depend on language, then our ability to assess altruistic behavior and thinking will naturally be strongly bound to language as well. This does not demonstrate that non-linguistic species do not have any sense of altruism or morality, but it does indicate that without language, and because of the apparent interdependence between language and morality, it is difficult for us to access and assess moral-like values in non-linguistic species.

2.2.2 *The Enforcement Function.*

Besides assessing and defining standards and tracking those who deviate from them, linguistically based social norms can also convert high cost altruistic behaviors into low cost behaviors (Sober and Wilson, 2000). These norms, connected to rewards and punishments would make self-sacrificial behavior less costly and transform altruism into a low-cost secondary behavior focused on enforcing norms. By way of example, norms and laws often come with rewards and punishments. The City of Los Angeles California uses a high occupancy vehicle (HOV) lane system. In designated lanes on the freeway, the rules only allow for cars with 2, sometimes 3 or more passengers. Because of the amount of traffic and one-driver cars in LA, the temptation to cheat on this system is great; however, the system is relatively successful for two reasons. First, language-based

enforcement system exists for punishing cheaters where freeway signs explain the penalty for violators.

The cost for the first violation is \$271, \$405 for the second, and \$675 for the third. With the large amount of traffic, the city could not police this system without language-based norms connected to costly punishments. Of course, some people cheat on the system, but because they know about the rule, which they can only understand through language, many people consider it too costly to risk cheating. Incidentally, the issue of social control through language will be a recurring theme throughout this thesis, and the HOV system is only one of thousands of other relevant examples. The system works through language; drivers understand the system through language; carpoolers report cheaters through language; and police contact and punish cheaters through language. Thus, language not only helps enforce this altruism, it lowers the cost of enforcement.

2.2.3 The Shared Representation Function.

Besides helping enforce norms, language enables us to communicate and share mental representations (Thierry, 2000). This helps us transmit a moral system to our conspecifics. On the other hand, though non-linguistic species have innate or learned abilities to recognize genetic relatives and behave in terms of inclusive fitness, they cannot linguistically share mental representations, and even if they have a mentally represented framework for morality, they cannot transmit it to each other for lack of a shared language. The above HOV lane scenario also exemplifies the importance of shared representations. Foreigners, with international driver's licenses but without knowledge of English would not understand the system, and they could violate it for this reason. Moreover, to take an extreme example, a circus bear that can ride a motorcycle would also fail to adhere to this system, not simply because of poor driving skills, but also because it lacks cognitive and linguistic skills. Hypothetically, trainers could teach a non-linguistic creature to adhere to the HOV rules, but intervention by trainers who know language invalidates the counter example.

Moreover, such training requires much individual attention and could not work on large groups. Shared representations also depend on "theory of mind," which most animals lack, and which interacts with language, values, and morality. That is, moral obligations almost always entail social relationships, and these relationship-based obligations require mutual "mind reading" to operate because people do not generally understand and respond to an obligation to another person without representing the other's internal state and how that moral/immoral action affects the internal state of the other. The moral injunction "Be kind to animals" can only work if we can imagine how animals feel. Contrast this with "Be kind to the furniture," which can only be understood metaphorically. Moreover, we use language not only to represent our moral obligations to others, but we also can employ it to ascertain and represent their internal states concerning our moral obligations to them. Language and theory of mind, therefore, greatly enhance social moral functioning; we negotiate with others, determining their state of mind by asking questions like "Is that okay?" "Do you mind?" "Did that upset you?" Later discussion will further demonstrate that morality involves maintaining social relations; however, this brief discussion shows that shared representation, a basic theory of mind, and language all enable morality to a degree that would be unimaginable without their support.

2.2.4 The Cooperation Function

Language not only helps us assess moral thinking, it helps people enforce altruism at low cost, and it supplies the shared representation system that promotes these functions. Furthermore, language helps speakers cooperate in order to foster altruism and morality (Knauff, 2000). That is, cooperation entails groupings, and symbolic language can help create groups and facilitate group cohesion and cooperation, thus playing an important role in group-oriented morality, which is an important form of cooperation, (Laland et al., 2000; Vine, 2000). Obviously many species form cooperative groups along kinship lines without the aid of language, but this still leaves us with interesting questions concerning the role of language and cooperation.

How does language help humans form cooperative groups? Does linguistic ability help us cooperate and exploit our environments more than if we did not have it? Regarding the first question, clearly a somewhat cooperative group could exist where all the individuals spoke different languages. However, their linguistic ability, at least initially, would be almost completely superfluous and would hardly aid cooperation. Since each individual would have the cognitive capacity to intuit what others might need, they could cooperate to a degree, but most of this cooperation would be possible because of the cognitive power of language and thinking. An individual would think to himself: "It looks like Giorgio needs help, and now he has looked at me with a "help me" expression, so I had better help him out."

Nevertheless, lack of language would severely limit cooperation. Imagine a scenario where the precise behaviors of 10 individuals are needed to rescue an elephant from sinking in quicksand. A leader shouts specific instructions about where and how each person needs to pull, push, apply levers, and hold onto ropes. In this case, language greatly enhances cooperation, and in a similar sense altruism and morality stand as a form of group cooperation that language significantly augments. Imagine a scenario where a group precisely defines rules for moral behaviors that will help them cooperate and coexist together. For example, the group defines the rules of possession and how to cooperatively punish people who violate those rules. Morality often involves cooperation, and most altruistic/moral rules have a cooperative focus. The injunction not to steal your neighbor's cow also represents a request to cooperate with your neighbor, and verbal directives to give to the common good promote cooperative and altruistic actions.

In addition, language also obviously serves as a cooperative aid for exploiting environments. A non-linguistic species can clearly remember the location of food and water, and to a degree they can signal about where such resources exist. However, except perhaps in the case of bee dancing or where humans have taught chimps and other animals abstract symbols for places, human language uniquely enables people to explain the location of resources through what is called displacement. That is, language enables its users to communicate about things not

in their immediate environment -- about the not-here and not-now. This clearly shows that language enables humans to adaptively cooperate and exploit environments in a way that would be impossible without it.

Besides this, history demonstrates that we express cooperative values in language. For example, many word-centered religions strongly emphasize that followers memorize cooperation-oriented moral precepts. The Deuteronomic tradition teaches followers to keep religious commandments in their hearts, *verbally teaching* them diligently to children, *talking* about them when sitting, walking, lying down, and rising up, and binding them as *written signs* on their hands, heads, doorposts, and gates. In such traditions, followers internalize a normative code that promotes cooperation -- with prohibitions against killing, stealing, adultery, and bearing false witness, and members can only internalize the cooperation code through language.

2.2.5 *The Group Fitness and Conflict Function*

Morality, descriptively speaking, can function for the benefit of the in-group at the expense of the out-group (Hartung, 1995). As previously mentioned, this concerns what a group might consider moral or prescriptive, what *they* think *they* ought to do, what *they* think is right and wrong. Such in-group moralities can brutally disadvantage or even terminate an out-group, or they can work in a way that benignly protects the in-group from out-group influences, with little or no detriment to the out-group. Even though we might consider such a "moral system" as ethically corrupt, it would still act as a normative and moral system within a group in a functional sense. These in-group moral codes stated in language illuminate inter-group conflict and the differential fitness of human groups in history, as Darwin himself suggested (1859).

Examples of this exist in virtually every culture. For instance, Japanese culture, which is well known for group-orientation, employed phrases in its history that express group solidarity for the purpose of defending itself against foreign influence, and incidentally, it is one of the few countries that successfully did not come under colonial rule. Under the Tokugawa Shogunate, Japan developed a

policy of *sakoku*, that is, "national isolation" and "exclusion of foreigners" to "close the country" (*sakoku-suru*) from foreign influence. Language helps in a situation like this because in large groups, many members will have never had contact with a foreign person; thus, the group easily marks outsiders with a negative symbol, and this serves as an efficient means to make distinctions and solidify the in-group as well.

Another historical example where a group used language in competition with outsiders comes from the Shibboleth story in the Hebrew book of Judges. Two Semitic tribes fight a war, and Gilead defeats Ephraim. Afterwards, Gilead captures the fords of the Jordan opposite Ephraim, and they guard a crossing place.

And it happened when any of the fugitives of Ephraim said, "Let me cross over," the men of Gilead would say to him, "Are you an Ephraimite?" If he said, "No," then they would say to him, "Say now, 'Shibboleth.' " But he said, "Sibboleth," for he could not pronounce it correctly. Then they seized him and slew him at the fords of the Jordan. Thus there fell at that time 42,000 of Ephraim. (Judges 12:5-6, New American Standard Version)

Here the word "Shibboleth" (meaning "ear of corn" or "stream in flood"), serves as a group marker, and if the potential outsider cannot pronounce the voiceless alveopalatal fricative (shi) and instead utters the voiceless alveolar fricative (si), he dies.

This simply and suitably exemplifies a uniquely contingent use of language to define group boundaries in a negative but clearly group selective sense. The linguistic password defines the group member, puts a skin around the group, and when insiders identify outsiders, they kill them, so that the in-group outcompetes the out-group in a life or death survival situation. That is, language helps define the group, resulting in at least a temporary reproductive fitness for the group that wins the battle, the lines of which in this case are drawn by language. Furthermore, the word "shibboleth" now refers to a behavior, practice, or an accent that identifies a person as an insider or outsider, and this also supports the point that human groups employ language for the benefit of the in-group at the expense of the out-group. Obviously, language, in this case, did not win the war, but it did serve as a marker, which the victorious side used to finalize their conquest. This reminds us that language can play a vital role to define groups and promote the kind of altruistic

behavior, such as risking one's life in battle that benefits the in-group at the expense of the out-group.

To summarize this section, (1) it is hard to *assess* moral norms without language, but language, for example, makes it easier to assess the moral behavior of deviants through gossip (Sober and Wilson, 2000). Moreover, (2) language provides a low cost means to *enforce* social control; for example, since punishment endangers the person who metes it out, it is much less costly to mark a deviant individual with a word than to punish him. In addition, (3) all this assessing, marking, and tracking behavior becomes more efficient through the *shared representation* system of abstract and symbolic language. Furthermore, (4) language significantly extends social control mechanisms and helps "allow groups to evolve into adaptive units" (Wilson, 2000), and this advances *cooperation*. Besides this, (5) linguistically based social controls also help facilitate *group fitness* in relation to *conflict problems* with other groups and environmental pressures.

2.3 The Degrees to which Human Morality is Possible Without Language

The above section summarizes how language uniquely serves as a mechanism that facilitates altruistic behavior or morality in human groups. Language provides an excellent means to assess both the conceptual and behavioral status of moral thinking and acting, and language enables us to mark symbolically what we consider immoral or moral. Additionally, linguistic diversity demonstrates that only insiders who share the same symbolic system can efficiently cooperate and function within a group's moral scheme, and languages clearly serve to mark group membership, thus facilitating group coherence in competition with other groups. These factors strengthen the case that humans cannot fully cooperate in a moral sense without language. However, this raises the question: how much of human morality is possible without language? How much do internally and mentally represented moral values require language, and how much do externally and socially represented morals depend on language? We can assess people's values and morals through surveys or by observing their behavior, but this does not show

us exactly how they depend on language for morality. This is the case even though we obviously depend on language for morality because we must use language when publicly teaching morality to and assessing it in others. For example, laws, moral codes, religious rules, ethical guidelines, all require language, and would not exist without it.

We may lack rigorous means to test our dependence on language for morality; however, we can design a thought experiment that exemplifies how morality could spread, or fail to spread, throughout a group without the use of any symbolic language. We place two individuals in an artificial setting for a 3 day period, where (1) we remove only one variable: all symbolic communication, including gesture, and (2) where Person A must cause Person B to understand who owns different piles of necessities that experimenters have placed in the room. Person A must also teach Person B that stealing these items is prohibited; however, A must teach B the rules without language. We reduce the morality to one concept of ownership and one principle: "do not steal." Even with this attempted reduction of variables, the scenario is still not simple. We have the categories, "I," "You," "resource-thing," "my-resource-thing," and "your-resource-thing." Subjects in this experiment still must understand, "I, You, It, My, and Mine" as well as the names of various things to be possessed. Moreover, they somehow need to understand each other's intentions about these categories without the use of language.

Person A acts as the initiator in this experiment. The experimenters explain the scenario to him -- that he must teach B the rules *without language*. Incidentally, it seems inevitable that we would have to permit some facial gestures. Person B thus only knows that she cannot communicate symbolically, but otherwise she knows nothing of the objectives of the experiment. (We must use language to explain the experiment to Person A, and to tell Person B that she cannot use language.) Then we place both subjects in a room with numerous piles of basic necessities, and as time goes on, Person A and B will become hungry, thirsty, hot, or cold (depending on how the experimenters control the room), and so they will begin to need things from the piles. A's goal is to help B understand who owns which pile *without using symbolic language*. Probably, the most he can do is use some kind of gentle

physical force, like pushing or blocking, since more dangerous forms would be forbidden. It seems clear that two people could come to understand without language who owns what, though perhaps their concept of ownership might not be the same as the one acquired through language because there would be no involvement of sense-relations between words such as 'own, belong, and property.'

However, once we start adding more people and more piles of things that each person will own separately, the communication of one rule becomes a complex problem. For example, one hundred subjects in the experiment would make almost impossible the process of trying to communicate ownership *without symbolic communication in a group of that size*. As many people start to need things in the piles of stuff, numerous variables would come into play. Some would hoard; others would share but discriminately so; some might steal, and still others -- because of the hoarders -- would be in want. Of course, this scenario is artificial and simplistic, but it shows how hard it is to communicate a moral system without symbolic language, *even when all of the subjects already have background knowledge about ownership and prohibition of stealing*. Moreover, most importantly, this shows that even a simplified moral system would fail when the group surpasses a certain size, a size that would make it impossible to employ a moral system without language, without resorting to pure physical force.

If we remove language and minimize paralinguistic communication in this scenario, we will find it difficult to communicate the moral rule, and this difficulty will increase as group size increases. The moral rule that exists in the head of Person A needs to be transmitted to the head of Person B in the simple one-on-one setting, and from Person A to 100 people in the more complex group scenario. Though we may find ways to communicate the rule in both settings, this thought experiment emphasizes the fact that language enhances the efficiency and effectiveness with which we can communicate the rule. Moreover, this does not diminish the power of paralinguistic communication to help us communicate empathically and morally, but rather this underscores how language enhances this process.

Similar to this thought experiment, Tattersall (1998) cites an analogous experiment done by the German-American psychologist Wolfgang Köhler, which tested the

ability of chimpanzees to pass on information about a complex learned task. The experiment vividly demonstrates the limitations imposed on the chimps by the absence of language.

[Köhler] taught a pair of chimpanzees how to obtain a food reward by tugging simultaneously on two ropes hanging a dozen yards apart -- too far apart for either chimpanzee to obtain the reward alone. Only a concerted action by both would suffice. Once the pair had got the idea, one individual was removed and replaced by another who had not been taught the same technique. The experienced chimpanzee clearly underwent fits of extreme frustration as the other failed to respond appropriately; he had a perfect intuitive understanding of the problem, but he could not carry out the task by himself and was simply unable to tell the other what to do to obtain the food. Of course, although he knew what to do, he had not arrived at the solution to the problem by rational thought; and it is vanishingly unlikely that he could have done so in the absence of the capacities that make language possible. For the abilities to arrive at the solution and to communicate it to others both stem from essentially the same mechanism. (Tattersall, 1998: 226-227)

Though the experimenter was able to teach the chimps without language, the chimps could not even pass on this knowledge without language, but we can assume that they would have been able to pass on the knowledge if they had linguistic ability. This shows how lack of language would seriously diminish the human capacity to cooperate, and this becomes even more problematic when we increase variables such as group size, the number of objects involved, and the number of behaviors we need to understand and evaluate. That is, communicating about physical objects is one thing. I can say to a child who does not understand the pulley system, "Pull the rope like daddy, and we'll get the food," and this will solve our problem. However, the situation becomes more difficult if we add more abstract concepts of ownership, stealing, good, and bad, or if we remove the situation in space, time, or reality from the subjects and objects involved. For example, "When we went to the Lake District on vacation, we should not have taken so many carrots from Peter's garden without asking; he will get upset if he finds out. I think we should apologize, and make restitution." In these ways, problems become even more complex and therefore even more dependent on language for their solution.

The issue of group size and communication relates to (Dunbar, 1996) whose research focuses on social grooming among primates as a mechanism for group bonding, and how linguistically-based gossip replaces grooming as group size increases. Dunbar suggests that as predation increases, primates experience two basic evolutionary responses: (1) to increase body size, and (2) to increase group size. At the same time, their primary evolutionary strategy is sociality, which causes them to spend 10-20% of their day in grooming each other as a means to maintain social bonds. However, the increase of group size places pressure on grooming; that is, group size tells us how much time primates need to spend in grooming, so the larger the group, the more they need to groom each other. When group size reaches a critical threshold where members must spend 30 percent of each day in grooming to maintain sociality, then it takes too much time, and more efficient grooming -- through language -- becomes necessary. Dunbar also shows a strong correlation between group size and neocortex size, suggesting that neocortex size not only predicts group size, but also that only large brained, linguistic species can maintain sociality as an evolutionary strategy. In short, larger groups and larger brains go together, and after groups reach a certain size, language becomes necessary to maintain sociality and the altruistic/moral bonds that enable groups to function well.

Though we will look at language as a group-bonding mechanism at a later point, at this stage, it is important to note how Dunbar's ideas relate to the above thought experiment. Physical-social grooming parallels how Person A shows Person B who owns what pile *without symbolic communication*. As group size increases, one-to-one communication becomes inefficient and costly regarding the time it takes to communicate the rule. Language simply increases efficiency by allowing Person A to communicate the rule to more people at the same time. In fact, communicating a rule through language is much more efficient than communicating social grooming through language. Grooming language -- oiling the gears of social relations -- requires some level of time and intimacy, and since it involves managing stress levels and maintaining positive affective distances between individuals, we do it best in a sub-group of a larger group.

However, moral language -- communicating the rule -- theoretically allows for more simple and efficient means to transmit information. People articulate rules and attach rewards and punishments to behaviors, and they accept a number of authorities to discipline offenders and reward those who follow the rules. In short, this thought experiment supports the claim, especially in light of Dunbar's research on language and group size, that in large groups without language a moral system is virtually impossible to transmit (or extremely inefficient and costly).

2.4 The Parallel Between Linguistic and Moral Development in Children

From the above discussion we can see that a robust relationship clearly exists between language and our ability to convey, comprehend, and comply with a moral system, as well as control a large group's response to a moral system. *Simply put, a large group cannot have a shared and complex moral system without language.* However, what about individuals? Though we will cover the issue about the moral development of feral or linguistically disabled humans in the next section, this discussion naturally raises the question about how moral development and linguistic development correlate with each other. According to Bhatia (2000) and Tappan (1997), little research exists that correlates moral and linguistic development in children. However, this area holds promise because initial findings appear to indicate an important correlation between moral and linguistic development.

On the theoretical side, Tappan argues from a Vygotskian perspective and considers language a "psychological tool [which] alters the entire flow and structure of mental functions" (Vygotsky, 1981). If so, as language mediates the moral side of social interaction, and as children internalize the moral ethos of this interaction, they develop mental functions that encompass moral thinking and behavior. Moreover, researchers will find this process hard to monitor empirically, because it is an "intrapersonal psychic process that manifests itself as inner speech" (Tappan, 1997: 94). However, in spite of the difficulty of analyzing the inner speech of developing children, in the Vygotskian view, linguistic and moral development parallel each other, and this shows some potential for giving us insight into these problems. For example, without asserting the innateness of moral ideas, Chapter 5 outlines

significant parallels between UG and human moral capacities. Moreover, it seems feasible, as Tappan claims, that language helps mediate moral action and that real moral ontogeny, development, and functioning cannot fully happen until children have the word-tools for categorizing and interpreting their actions and the actions of others, as good, bad, right, or wrong.

This relates to Hanfling (2003) who proposes that in learning words like "steal," children generally learn that stealing is wrong, thus implying that children learn moral concepts by acquiring the meaning of morally laden words.

Given [the environmental conditions required for language learning], it follows as a matter of human nature that the child will pick up the use of language, and this includes the language of right and wrong, vice and virtue... Knowing the difference between right and wrong, between virtue and vice, is nothing other than knowing what acts are describable as 'right', 'wrong', 'virtuous' etc.: it is a matter of being able to apply these words correctly. (Hanfling, 2003: 25-26)

For this to be possible we may need the ability to process the concept of 'wrongness' at some innate level; nevertheless, a couple of interesting characteristics about language and morality emerge from Hanfling's claim. The first concerns how we distinguish between right and wrong and how we act on what is considered right, good, altruistic, or moral. Moreover, other than truly accidental and random acts of altruism, it seems consistent with experience that knowing something about right or good precedes acting on it. The second characteristic concerns the argument that we can simply change the value attached to the word 'steal.' For example, "I hereby claim that stealing and greed are good." However, like other justifications of stealing, this misappropriates or misinterprets the word. That is, this is not a correct use of the word 'steal.' Hence, 'to steal,' means to wrongly take something that belongs to another person; thus, in basic terms, learning the word implies learning the value attached to it, even though non-stealing behavior does not necessarily come from learning the word. These two characteristics strengthen the case that moral development feeds off of language development, and this supports a main thesis of this project that morality cannot exist without language.

For example, simply exchange the roles in the above thought experiment from Person A and Person B -- to Parent A and Offspring B. How is the parent hindered in raising a child without language? If the child were deaf, blind, *and* without Braille ability, how would the parent have access to the child's mind in order to dialog about categories of good, bad, right, and wrong? (We will look at Helen Keller's case later on). The parent probably could physically intervene with the child's behavior and instruct the child how to morally interact with others and her environment. This does not mean that the child could not learn by herself, and a non-linguistic child who receives affection and regard from caregivers should be able to show affection and regard to others. However, such a child would experience great difficulty in learning from others. Furthermore, even if the child has sight and can learn from observing others, if she cannot use sign or spoken language, then neither can she experience any of the moral development that comes through language. Moreover, if we say that human moral development can *completely* take place in the absence of language, we also render practically meaningless the reality that much language usage focuses on moral development, such as religious language or caregiver-child interaction about right and wrong. In short, the absence of language should make moral development more difficult, and the presence of language should make it easier.

Incidentally, the idea that language facilitates moral ontogeny substantiates a notion to be discussed later, that language may influence moral perception. If language proves instrumental to the moral growth of children, then moral development embraces elements unique to historical, social, and cultural particulars. People generally assume that global variation in moral and societal values stems from cultural differences between ethno-linguistic groups, and research validates this assumption (Hall and Hall, 1977). However, if language robustly, or even moderately serves as a mechanism for processing human moral ontogeny, then cultural and moral *differences* may emerge in individuals through a culturally *and* linguistically guided process. This is not to say something trivial, for example, that differences in syntax between English and French create putative differences in moral character of the cultures of England and France. However, a

linguistically influenced moral ontogeny suggests that linguistic differences might play a role in shaping moral development *differently* across cultures.

To clarify, this is not to say simplistically that language creates the cultural differences we see in the world today. However, sociolinguistic factors like register, such as Japanese honorifics, or semantic emphases based in cultural values, such as Japanese words emphasizing harmony (*wa wo tamotsu*), 'to maintain harmony,' or the numerous Japanese words emphasizing 'inside/outside' distinctions (*uchi/soto*) may cause children to pay attention to some values more than others. For example, these linguistic tendencies might cause children to focus more on harmony or insider/outsider issues than children whose languages do not commonly collocate phrases with emphasis on these issues as much. Chapter 6 deals with these issues in depth. However, in short, since common collocations and language usage make some social features more salient in one language than another, this underscores how linguistic differences might differentially affect moral ontogeny across cultures.

Returning to the question of how language may guide moral development, Snow states (1990: 113) (commenting on the work of Dunn) (1987) that "an enormous amount of talking about moral spheres occurs in natural contexts among very young children." This suggests that linguistic analysis of children's talk about morality could prove interesting. Snow herself approached the problem using a lexical concordance; she analyzed 51 hours of caregiver-child transcripts in 30 second to 15-minute segments taken over irregular but frequent intervals over a period of 3.5 years. The data consisted mainly of conversations between a father and son, during a period when the son was 2.5 years old until he was 6. Snow looked for occurrences of the words *awful*, *bad*, *good*, *mean*, *naughty*, *nasty*, *nice*, *right*, *ugly*, and *wrong*, focusing on when they were explicitly used in a moral or deontic sense. She summarizes one of the most interesting results of this research that shows a correlation between an increase in linguistic and moral complexity:

The frequency data... suggests that whereas talking about "bad things" dominates moral discussions at two-and-a-half, the emphasis shifts to "good" at three-and-a-half, and increasingly at ages four-and-a-half and six to "should and "shouldn't." Perhaps this is not surprising, for it parallels an increasing

sophistication in the kinds of moral reasons revealed in the discussions. (1990: 117)

Regarding the centrality of language in moral ontogeny, Snow (1990: 120) describes the process as a "negotiation of the meaning of words and of acts." Obviously, negotiating about *words* and actions comprises a linguistic process. Moreover, Snow says that in this negotiation, parents also "exercise the right to define the child's behavior as good or bad and to relate their goodness or badness to the intentions behind and the consequences of the acts" (1990: 120). Of course, *defining* behavior also involves a lexicological and thus linguistic process, where parents describe and categorize behavior as good or bad. Obviously, the effectiveness of this process depends on many factors, such as a trusting relationship between the caregiver and the child; however, the important point here confirms that the process of moral ontogeny cannot fully happen without language.

We should not over-generalize from one case study in Snow's work, but it does show the promise of a theory of *linguistically led moral ontogeny*. As Snow states:

Throughout this period, and into adolescence and adulthood, we must recognize the role of *narrative* in making both *words* and actions morally *meaningful*. Why is an act good or bad, after all? Because of the *story* that is told about it. Most acts are morally neutral if not viewed in the context of the intentions that led to them and/or the consequences that derived from them. (1990: 121, emphasis mine)

Moreover, this meaningful narrative about intentions, behaviors, and their consequences reveals a meta-narrative that moral ontogeny requires the scaffolding and support of language. Clearly moral ontogeny comprises a complex process, involving many psychosocial factors; however, the argument here suggests that this process is not simply complex, but it may also be non-reducible -- in the sense that we cannot remove language from it. Thus, if we remove language from the developmental process, then we may severely stunt or limit moral ontogeny.

2.5 Moral Ontogeny: Broader Considerations

Nevertheless, we must not overestimate our understanding of this complicated bio-neuro-psycho-sociological process. Moreover, since this study focuses primarily on

the linguistic aspects of moral development, it is beyond our scope to fully consider every possible aspect of moral ontogeny. However, to provide sufficient background, we must consider aspects of development that transcend linguistic considerations. Psychiatrist Allan Schore (1994), writing from a psychoanalytic perspective, summarizes a wealth of information on child development, including 2300 references, and he provides numerous persuasive arguments that suggest language gives essential support to moral ontogeny. In sum, he claims that the dyadic relationship between child and caregiver supplies a rich physical, emotional, visual, and verbal socialization process that plays a crucial role in the development of a well-adjusted child, including the child's moral development. In a typically pithy summary of data, Schore says:

This dyadic psychoneurobiological mechanism ontogenetically sculpts the enduring temperamental features of the child's emerging personality. It also generates emotional biases that determine whether an individual's or culture's processing of a particular affect will generate a conscious or unconscious emotional state. (Schore, 1994: 282)

Though Schore does not mention language here, in any caregiver-child relationship, language will play a central role in a normal process that shapes a child's temperament. Obviously, a person's enduring temperament can deeply affect their moral self in how they treat others, or how they experience empathy for others. Schore also implies that this relational ontogenetic mechanism will produce cultural differences in the way people affectively process information, and this also relates to the idea that moral ontogeny and perception vary across cultures. That is, as language influences development, it may cause cross-cultural variation in how individuals perceive emotional states.

Regarding empathy and the beginnings of moral development in children, Schore claims the ability to respond with distress to the distress of others requires a period of relational social referencing with a caregiver who is emotionally engaged with the child. Thus, in the neurology of moral development, "The right hemispheric orbitofrontal affect regulator performs functions that are essential to the adaptive moral functioning of the individual" (1994: 353). Moreover, this right hemisphere orientation results from the visual-facial stimuli that children receive from their

primary caregivers. This is interesting for this linguistic discussion because affect regulation develops in the right hemisphere, but in most people language localizes in the left. This would be a problem for any simplistic claim for the linguistic basis of moral development and functioning in humans. Therefore, any theory about the relation between morality and language must be sufficiently complex and nuanced. Schore elaborates:

At 18 months the child first exhibits moral prosocial altruistic behavior in the form of comforting -- regulating the negative affect of a -- distressed other. This other-oriented empathy is attained when the child is capable of reading her own as well as the other's negative internal state. Such a capacity necessitates the operation of an internalized shame regulator. These formulations suggest that moral development begins earlier than previously thought, and is much more of a visual than a verbal process. (Schore, 1994: 354)

Schore asserts the *beginning* of the child's moral development is visual -- based on feedback primarily from the caregiver's facial communication. Hence, this aspect of the genesis of morality in empathy is not verbal or linguistic in nature, but the communication occurs, primarily through "abstract mental images of faces" (1994: 186) and facially expressed shame induction, which come through the caregiver's disgusted face. The process is visual and imagistic and tends to influence the right hemisphere in a highly affective way. Thus, a caregiver helps "shape the child's developing moral capacities" (Schore, 1994: 354) and even the neurological structure that mediates moral functioning. Moreover, Schore emphatically states that the orbitofrontal cortex (OFC) contributes vitally to empathic and moral behaviors, and as other research broadly supports, socio-pathologies appear to emanate from damage to the OFC (Anderson et al., 1999; Dolan, 1999; Blair and Cipolotti, 2000; Brower and Price, 2001; Moll et al., 2002).

This data helps present a full and balanced discussion on language and morality because it demonstrates the complexity of the process of moral development, the interaction of essential variables, such as facial imagistic stimuli, necessary physical affection, as well as deontic linguistic input. Schore explains on these social factors:

For humans, as for all other species, the most salient aspects of the environment are located in not so much the physical as in the social context, the realm of

interactions between one individual and another. The transactions within human relationships are both verbal and nonverbal, and thereby contain both cognitive and emotional elements. (Schorer, 1998: 337)

Schorer suggests that as the caregiver holistically communicates with the child, this communication facilitates a biological, mental, and social response that influences the child's ability to evaluate social situations and thus respond in socially adaptive ways. The significant role of imagistic facial stimuli does not negate the role language plays in moral ontogeny, but rather indicates that verbal factors will stand as one aspect of an integrated and essential set of variables that help socialize developing humans.

In addition, as a caregiver nonverbally and symbolically expresses facial approval and disgust to the child, and as this provides a socio-affective, neurological, and non-linguistic basis for empathy and morality, Schorer nevertheless correlates *moral* and *language* development. First, there is general agreement that single and two word utterances develop before syntactic speech (Wray, 2002). We may also consider these words as holophrases in which "Mama!" could mean, "Mommy, come here" or "Here comes Mommy." Accordingly, Schorer asserts, based on considerable literature review, that single words such as "no," "good," and "bad" undergo right hemisphere lateralization. Moreover, these words are affectively charged concepts that play an important role in self-regulation through adulthood. Furthermore, considerable consensus exists that emotional experience (Van Lancker, 1997) and affective prosody (Ross et al., 1997) find greater representation in the right hemisphere.

Besides this, Schorer stresses that the child's own name undergoes right hemispherical processing, and it is clear that caregivers use the child's name affectively for managing stress and comfort. A child's name requires language, stands central to her identity, and enables her to conceptualize herself symbolically, and without a distinct concept of self, it seems unlikely that self-regulation, empathy for others, and thus morality would fully emerge. In short, though children experience the imagistic and visual side of moral development in the right hemisphere, they experience linguistic and verbal moral development in the right hemisphere as well.

Consequently, this crucial developmental period is "influenced by dyadic affectively-laden verbal transactions" (Schorer, 1994: 489) that support language learning. Moreover, "the mother's socializing communications, used to inhibit and direct the toddler's actions, are expressed in prosodic patterns that highlight affectively salient events." Caregivers articulate these patterns with "face engagements" that imprint onto the "developing orbitofrontal-superior circuitry that is involved in auditory-visual cross-modal learning" (1994: 489). This is the rich environment in which the child generates single word speech that is high in affect, self and other valuation, and therefore moral overtone. The combination of these theoretical and empirical considerations causes Schorer to comment:

In keeping with the principle that the right hemisphere develops before the left, the early language that appears in the second year represents the output of the right hemisphere. The socioaffectively-driven expansion of language in the service of right hemispheric cognition thus mediates the development of self-regulation through private speech, that is, thought. (Schorer, 1994: 489)

Thus, in this complex scenario that Schorer describes in the caregiver-child relationship, the developing child will learn to construct and represent her affective interaction with the world in autobiographical and imaginative dialogs. If the child is not deprived of care, she internalizes the caregiver's comforting verbalizations and employs them through self-caring evocative memory. Schorer even says that if the caregiver or circumstances deprive the child of this process, she may develop the regulatory disturbance of alexithymia (Greek for "no-speech-soul"), that is, an inability to appropriately or non-violently recognize and express emotions. Moreover, as the affectively charged words of "good" and "bad," and the child's name imprint on the child, she learns to regulate herself in relation to (1) shame in this dialogical socialization and (2) the need to attune to the caregiver's injunctions to "be good." In short, psycho-moral-socialization appears to consist of an inexorably intertwined matrix of factors that influence how the child develops her moral self. Most importantly for this discussion, language is one of the essential factors for normal socialization that produces a well-adjusted individual who can function in a community where people express adaptive sociality through appropriate levels of regard for others and acceptable moral behavior.

2.6 Language and Morality for Feral or Abused Children

In addition to questions about how moral ontogeny depends on the above factors, this discussion leads us to consider how language relates to morality for developmentally disadvantaged or disabled children. Such children generally come out of a defective and unhealthy socialization process -- either as neglected, abused, or brain-injured children. Because feral and sociopathic children are often the products of extreme neglect and/or abuse, there are serious ethical concerns involved in this area of research. Moreover, most study concerning these people is clinical and related to their rescue, recovery, or readjustment to normal social relationships after being "found," or as the following discussion suggests "captured" may be a better term for sociopaths whom society institutionalizes for criminal behavior.

Regarding sociopaths, (Lykken, 2000: 559) discusses what he calls "feral youngsters," children who experienced an impoverished socialization. Lykken calls these individuals "feral" sociopaths because they are genetically normal (as opposed to psychopaths), but their behavior is extremely violent, criminal, and often irremediable. Lykken's views are controversial especially because he suggests parental licensure as a cure for this problem. He thus aims at the basic socialization process -- requiring parents to study, prepare for parenting, and obtain a license before having children. For this discussion, we need not critique Lykken's suggestion, but rather consider his premise that sociopathic behavior stems from a maladaptive socialization process. In short, Lykken thinks that a broken socialization mechanism is a primary cause of sociopathic behavior resulting in high crime statistics in the US. This does not prove a counterclaim that if we socialize people "properly" they will become "moral" citizens; however, the important point here demonstrates that negative social patterns may indicate a flawed socialization process.

In flawed socialization, if the affective factors of trust and parental investment are inadequate, then what does language do or not do to help socialize feral people? Clearly, most sociopaths exhibit full human language capacity along with their bad

behavior; however, this does not negate the role language plays in human moral development. That is, the claim that morality cannot exist without language does not entail that if language exists, therefore morality exists also. In fact, skillful speech in the mouth of a sociopath will make him much worse. Therefore, as already stated, language stands as one part of a non-reductive matrix of factors that forms morality in people, and regarding sociopaths, another factor must have been missing so that they could not develop normally. Thus, language is necessary but not sufficient for moral development.

Another way to view this issue considers how verbal abuse or extreme neglect would affect the moral ontogeny of a child. How would a child develop morally if, for example, his parents or caregivers always spoke to him maliciously? How would children fare morally if we never spoke to them at all, or what if we exposed them to a care-giving narrative that placed positive value on antipathy towards others, on categorizing stealing, cheating, and brutality as good, and generosity, honesty, and kindness as bad? This would be a kind of morality (anti-morality) as it prescribes values. Nevertheless, it is likely that such a warped socialization process would produce sociopaths even though investigating it would be difficult and ethically impossible. For moral socialization, the existence and quantity of linguistic input appears far less important than its nature and quality.

In addition, when sociopathic parents raise children, what kind of moral narrative and dialog takes place? What are the salient characteristics of such a maladaptive socialization process? Would it be possible to quantify the linguistic-moral-ontogenetic data in a parent-child relationship, as (Snow, 1990) does in the discussion mentioned above, and if it were possible and ethical to obtain and evaluate such data, then what would that data reveal? I conjecture that it would reveal a lower frequency of the behavioral factors that Schore (1994) mentions, including morally oriented narrative -- that defines and categorizes behavior as good or bad. Moreover, though some parents might carefully teach their children how to be a skillful criminal, murderer, or sexual predator, they would more likely leave their children to *develop in a vacuum*, and thus develop morally on their own. If they treat their children cruelly, they would thus teach sociopathic behavior to

their children by example and inflict serious emotional damage on them as well. Moreover, other key factors in a moral matrix, such as physical, affective, and social care would likely be qualitatively and quantitatively deficient, which would greatly dilute many -- if not all -- the factors that help people develop a moral outlook that enables them to function within a particular group.

The above outlines some of the ways poor socialization might increase the possibility of sociopathic behavior in people; however, Lykken also describes the strong interrelationship of all these factors in the following discussion.

How do most children avoid becoming social misfits? Probably in much the same way as the young of other social mammals learn the rules of their communities, through the monitoring and example of their elders. In southern Africa, the white rhinos were being murdered, not by poachers, but by young male elephants who were orphaned by culling operations in the Kruger National Park (Lemonick, 1997). The adults of the matriarchal herds were shot and the baby elephants transported to other parks where they grew up without the normal years of parental supervision -- and they grew up to be dangerous outlaws. Our species ranks between the elephants and the great apes, toward the low end, and the ants and hymenoptera, at the high end of the continuum of socialization. We are born with the capacity to develop a monitoring conscience that works to inhibit rule breaking. We can learn to feel empathy for our fellow creatures and to take satisfaction in acts of altruism. Most of us develop a sense of responsibility to our families and our community, and desire to pull our own weight in the group effort for survival. We may be the only species with a strong, clearly differentiated self-concept so that we are motivated to emulate people whom we admire in order to feel good about ourselves.

Unlike the hard-wired proclivities of the social insects, however, these prosocial inclinations do not emerge in us as well-formed instincts but, like our inborn capacity for language, they require to be elicited, shaped, and reinforced by our interactions with other, older humans during our early development. Our poor success in rehabilitating persons who have reached young adulthood still inadequately socialized suggests that, again like our language capacity, there may be a critical period for socialization. Unless it is evoked, sculpted, and made habitual in childhood, our human talent for socialization may wither and never develop. (Lykken, 2000: 567)

Lykken provides a striking parallel between outlaw elephants and immoral humans, both products of a shattered socialization process. This relates to de Waal's (1996) substrate morality in non-human species and what happens to these species when the socialization process breaks down. What can we learn from

animal socialization regarding substrate morality? For this discussion, we can tentatively say that a stunted socialization may produce asocial or sociopathic behavior in most social species -- even those without language.

This brings to mind the story of Helen Keller whose socialization was greatly enhanced by her breakthrough of learning to communicate with language. This is not to state the non sequitur that disabled people without language are subhuman, immoral, or without rights; but this simply acknowledges the extreme difficulty of socializing a person with whom we cannot communicate linguistically. Moreover, a non-linguistic person may still receive natural care and kindness throughout development and thus become a gentle and empathic individual; however, this differs from a fully socialized person who is able to effectively function in a society that employs language at the center of its activities. All in all, these factors reinforce the fundamental link between language and human socialization.

2.7 Language Innateness and Moral Ontogeny

Lykken also compares human language and moral capacities. That is, as we shape and socialize our inborn language capacity, we also must socialize our moral development, and as a critical period exists for acquiring language, one also appears to exist for moral socialization. Lykken's correlation between innateness, hard wiring, a critical period, socialization, language, and morality requires a more rigorous comparison, and Chapter Five explores parallels and non-parallels between Universal Grammar (UG) and morality. However, the question here concerns language innateness and moral ontogeny. If language helps socialize people morally, and if it is extremely difficult for sociopaths to acquire a socially adaptive morality after a certain age, a correlation may exist between moral and language acquisition. A correlation does not necessarily imply a causal relationship, but at least this suggests that maturation constrains both linguistic and moral ability. Moreover, as already discussed, if we remove language from the socialization process during the critical period, then we will produce a partially feral child, and such a feral child will lack any knowledge of morals that come through language.

In addition, moral ability like linguistic ability, at least partly appears to stand in dissociation from general intelligence. Schore (1994) cites numerous studies that link morality to the orbitofrontal cortex (OFC), which operates in moral decisions and moral appraisal. Subjects who have experienced damage to the OFC suffer serious moral dysfunction and dyscontrol (Anderson et al., 1999; Brower and Price, 2001; Moll et al., 2002), and this is similar to the affect left hemisphere damage has on language ability. Patients who underwent adult-acquired damage to the OFC could reason well about hypothetical moral dilemmas, but they could not apply the moral knowledge in real life. However, subjects that endured early-onset damage made excellent recoveries, but later displayed severe moral dysfunction and dyscontrol, and besides the real-life problems, they showed serious deficit in their moral reasoning abilities (Anderson et al., 1999).

The temporary recovery of the early-onset victims is puzzling. However, it may imply the following. The moral reasoning that is unimpaired in late-OFC-damage is part of linguistic, not centrally moral behavior. Thus, two etiologies exist for amoral behavior. First, late-OFC-damage may affect knowledge of 'ought' accompanied by the ability for moral reasoning in which the victim lacks a control system to apply this knowledge to behavior. Second, early onset damage causes a lack of knowledge of 'ought' where the victim lacks moral knowledge to apply to behavior. This implies that theoretical, linguistically expressed knowledge of rules about 'ought' is located in a different place from the inhibition mechanisms that apply that knowledge to behavior.

Moreover, environmental explanations for this maladaptive behavior eluded the researchers; the children grew up in stable middle class environments and had siblings who were socially fit. This seems to isolate OFC damage, not dysfunctional socialization, as the primary cause; therefore, evidence suggests a critical period for moral acquisition. That is, those with late OFC damage experienced only behavioral dysfunction; however, those with early onset damage to the OFC experienced both moral appraisal and moral behavior dysfunction. Since early onset injury to the OFC damages both moral appraisal and behavior but late onset injury only affects behavior, this implies that the OFC serves both

behavior and appraisal in early life. However, the main point here remains that moral and linguistic capacities are similar because they experience modular organization in different parts of the brain, and both require extensive socialization.

This discussion also raises another complex problem. If different brain modules control language development and moral development, or if these abilities localize in different parts of the brain, then should we expect some dissociation between language and morality? The answer is clear; a neurological dissociation between language and morality still does not suggest that language does not play an essential part in human moral development. Damage to the OFC that affected moral function, judgment, and behavior does not affect linguistic ability, and damage to Broca's and Wernicke's areas does not affect moral judgment or behavior. Nonetheless, damage to Wernicke's area affecting linguistic comprehension could make it difficult for subjects to understand and then appraise moral dilemmas, but this difficulty in appraisal would stem from an indirect cause, and not a direct cause affecting the brain locale implicated in moral function. The dissociation of morality and language in the brain reminds us that language is not morality, nor is morality language, but the dissociation does not discount the strong interplay between the two.

2.8 Apparent Altruistic or Moral Behavior in Non-Human Species

Having looked at morality from a developmental perspective and how the presence or absence of language influences moral systems, our attention now turns to morality in nonhuman animals. Darwin wrote in Chapter Four of the *Descent of Man*:

I fully subscribe to the judgment of those writers who maintain that of all the differences between man and the lower animals, the moral sense or conscience is by far the most important. (Darwin, 1882: 97)

If the moral sense stands as a most important difference between man and animals, then it potentially provides a salient means of comparison that can offer additional insights in this discussion on the how language interacts with morality.

Here again it helps to clarify terms. As mentioned before, the general term "altruism" refers to any kind of other oriented behavior, including reciprocal altruism and kin selective altruism. "Genuine altruism" and similar terms refer to non-kin-directed, non-reciprocal altruism, which incur a fitness cost but with less or no return benefit to the altruist. "Morality" refers to how individuals or groups define or categorize behaviors as good or bad and how people *ought* or *must* do good and not bad. This clarification is important for two reasons: (1) because this will help us distinguish between the moral senses of animals and humans, and (2) because orthodox Darwinism claims that genuine altruism does not exist in humans, in fact that it does not adaptively exist anywhere in the natural world at all.

This second clarification is important because despite the mainstream consensus behind Selfish Gene Theory (SGT) and the claim that genuine altruism does not exist in biological nature, conceptual quandaries and natural facts nevertheless challenge the status quo. Thus, many prominent scholars have criticized SGT because of these problems, (Midgley, 1991; Lewontin, 1993; Rose, 1997; Dover, 2001; Gould, 2001). For example, Primatologist Frans de Waal says, (1996: 1) "For one thing, inasmuch as moral rule represents the power of the community over the individual, it poses a profound challenge to evolutionary theory." That is, some kinds of altruism and morality may exist in nature -- not just in humans -- that SGT either fails to explain or simply ignores. To understand how this is the case, we need to see what proponents of SGT have said about generous and costly altruism. The following quotations briefly expound on the SGT view of altruism, and I have underlined key phrases while leaving italicized phrases unedited.

Be warned that if you wish, as I do, to build a society in which individuals cooperate generously and unselfishly towards a common good, you can expect little help from biological nature. Let us try to teach generosity and altruism because we are born selfish. (Dawkins, 1976: 3)

If there is a human moral to be drawn, it is that we must *teach* our children altruism, for we cannot expect it to be part of their biological nature. (Dawkins, 1976: 150)

We can even discuss ways of deliberately cultivating and nurturing a pure, disinterested altruism -- something that has no place in nature, something that

has never existed before in the whole history of the world. (Dawkins, 1976: 215).

Let us understand, once and for all that the ethical progress of society depends, not on imitating the cosmic process, still less in running away from it, but in combating it. (Huxley, 1893)

A selfish act is one from which the actor gains a benefit while it inflicts a cost on the recipient. Natural selection operating on the actor favors selfish acts, but the potential recipient is selected to avoid receiving the costs. We can speak of a co-evolutionary struggle between the tendency to perform selfish acts and the tendency to avoid being harmed by the selfish acts of others. (Trivers, 1985: 52)

To be blunt, my Darwinian ethics says that substantive morality is a kind of illusion put in place by our genes, in order to make us good social cooperators. (Ruse, 2002: 163)

Sociobiology suggests that the road to cooperation is paved by selfishness. However, the rub is that not only is the pavement concretely selfish but also the destination itself -- the only terminus in the journey of life -- is self-interest. (Post et al., 2002: 215)

Nevertheless, in spite of these claims, if some form of real altruism exists, SGT does not fully explain altruism and morality. Either these scholars are stating the explanatory limits of SGT (which is not their intention), or they are failing to include the possibility of real altruism and morality in their theory of nature. As de Waal succinctly puts it: "By thus locating a morality outside nature, the scientists have absolved themselves from trying to fit it into their evolutionary perspective" (1996: 16). SGT proponents may reply by saying that the theories of kin selection or inclusive fitness (Hamilton, 1964) and reciprocal altruism (Trivers, 1971) completely explain all examples of altruism. However, though these theories account for much altruism as an expression of genetic self-interest, the problem remains that expressions of genuine altruism that transcend the selfish-gene model may exist, not just in humans but also in other animals.

Moreover, the existence of genuine altruism in lower animals is especially interesting for our discussion about the nature of morality and language for two reasons. First, for instance, if animals exhibit altruism that might benefit the group at the expense of the individual or his genes, then this would stand as precedent for the existence of an even greater altruism in humans. This coincides with Bekoff's

research, which suggests that animal nature "isn't always ruthlessly selfish and competitive" (2002: 34). Second, to expand on this, if animals with minimal or no culture and with minimal or no linguistic ability can express altruism, then humans with the benefits of culture and language will have even greater resources to facilitate, support, and extend altruism. Moreover, we can test this claim while still acknowledging that humans can excel at egoism and moral evil, egoism and moral evil that linguistic ability enhances, for example, through the power to deceive. Hence, though language has the power to increase altruism and morality, this capability is also a double-edged sword, cutting in both selfish and altruistic ways; nevertheless, looking at animal altruism may help us better understand human altruism and morality, especially in the light of language.

2.9 Animals, Language, Altruism, and Morality

Regarding animals, altruism, and morality, we can focus on some basic questions.

- What kinds of altruistic or moral-like behavior exist in non-human animals?
- Do animals express altruistic behavior to genetically unrelated conspecifics?
- What is the difference between this kind of behavior in wild animals compared with semi-wild animals in zoos or research observation centers?
- In these trained or domesticated animals, does evidence exist that language influences their moral behavior, or does that behavior remain unaffected by the presence of humans and language?
- How is human morality different from animal morality, and what role does language play in creating these differences?

In short, answers to these questions should provide us with important background information and empirical data on how language affects altruism and morality.

First, regarding the moral-like behavior of animals, numerous researchers are developing an interesting literature on the topic. Preston and de Waal (2002) describe perception-action processes related to empathy that exist in many species, and the authors supply a theoretical and empirical framework for researching empathy in nonhuman species. This framework provides a view that "can advance

our evolutionary understanding of empathy beyond inclusive fitness and reciprocal altruism and can explain different levels of empathy across individuals, species, stages of development, and situations"(2002: 1). Hence, empathic altruism may differ from SGT explanations of altruism, and this echoes what Darwin said in the *Descent of Man*.

The social instincts lead an animal to take pleasure in the society of its fellows, to feel a certain amount of sympathy with them... Sympathy forms an essential part of the social instinct, and is indeed its foundation-stone... Many animals, however, certainly sympathize with each other's distress or danger. (Darwin, 1882: 99)

Along these lines, Preston and de Waal argue that empathy forms an evolutionary basis for altruism in many animal species, and thus empathy may provide a foundational element for the development of human morality as well. Incidentally, the terms *empathy* and *sympathy*, though distinct, are used interchangeably here to refer to how an animal can empathize with the physical and emotional distress of a conspecific. This argument, therefore, simply increases the rigor of Darwin's original contention that sympathy/empathy exist as a social instinct basic to the evolution of human morality.

On the theoretical side, they suggest that the "perception-action organization" evolved in the nervous system as an antecedent to empathy. Organisms thus adaptively respond to perceptions, becoming attuned to the caregiver, and developing empathy through the process. On the empirical side, the authors cite numerous experiments where animals such as rats, rhesus monkeys, and human infants exhibit empathic behavior, distress at the distress of others, and therefore this shows "that individuals of many species are distressed by the distress of a conspecific and will act to terminate the object's distress, even incurring risk to themselves" (Preston et al., 2002: 3). For example, regarding a monkey experiment (Masserman et al., 1964), the authors comment:

These monkeys were literally starving themselves to prevent the shock to the conspecific. Starvation was induced more by visual than auditory cues, was more likely in animals that had experienced shock themselves, and was enhanced by familiarity with the shocked individual. (Preston et al., 2002: 1)

Examples like this suggest that *empathic altruism* may exist as an explanatory tool beyond inclusive fitness and reciprocal altruism that can help us understand the biological basis for empathy, altruism, and even morality. More importantly, it raises questions about the differing degrees of empathy that various species can feel towards kin and non-kin, and it shows that animals might direct empathic altruism to non-kin, to non-similar species, and even to non-reciprocators. This is the case because if empathic altruism stands as distinct from reciprocal altruism and inclusive fitness, then it could also transcend the limits of selfishly genetic altruistic behaviors.

Preston and de Waal discuss cognitive empathy in their model, which expresses itself in the so-called markers of mind, such as perspective taking, mirror self-recognition, deception, and tool use. For example, studies of cognitive empathy demonstrate that apes show extended helping and consoling behavior (de Waal, 1996). Though it is not clear how visceral, basic, and automatic forms of empathy relate to the cognitive forms, basic empathy together with cognitive empathy probably can extend helping and consoling behavior.

Basic empathy combines with cognitive empathy; for example, humans can feel emotional distress at the distress of another and they can think about and act on ways to relieve that distress. These two abilities appear distinct; for example, as cited above, a rat that expresses distress at the distress of another rat does not experience this emotion to the same extent as a primate. Therefore, monkeys or chimps because of their greater intelligence must experience empathy at a greater emotional and cognitive level than less intelligent animals, even though this greater empathic ability still exists without language. This does not mean that intelligence correlates directly with empathy, for an autistic person may show high levels of intelligence without empathy for others. Nevertheless, as stated before, some cognitive abilities must correlate with empathy. In short, this brief section shows that empathy in nonhuman species parallels human social and moral behavior and suggests that symbolic language can extend empathy and reciprocity, making language-enhanced empathy basic to human moral codes.

2.10 Symbolic Language, Empathy, Reciprocity and Moral Codes

Most importantly, Preston and de Waal's research implies that language can ratchet up empathy in significant ways. In discussing the traits of working memory, planning, and inhibition, they state the following:

With working memory, individuals can hold information in mind and manipulate this information to predict, compare possible outcomes, and decide on an appropriate course of action. Working memory also increases imaginative processes that allow individuals to evoke empathic processes in the absence of the object. (Preston et al., 2002: 20)

Though the focus here is on working memory, this clearly links up with language since language in combination with memory enhances abilities for predicting, comparing outcomes, and decision-making. For instance, argument structure makes possible these empathic words and their thoughts "that could be me" or "they could be us." Therefore, with language we can think: "That person sleeping outside in the cold could be me, and if I don't help him, he might get sick, so maybe I should give him my old coat." Moreover, if a person feels tempted to do something selfish and illegal, then he can 'talk' himself out of it by verbally reminding himself of the potential outcomes. On the other hand, we can also talk ourselves out of an empathic response. "I had better not help that guy; it might be dangerous." Though the empathic response may be visceral, language can help decrease or increase this response in individuals or in larger groups. Altruists may often use language to ratchet up the empathic response of a group in order to increase helping behavior, for example, when a leader tells a group to donate to charity.

Besides helping increase and extend the empathic response to the distress *of others*, language may also help inhibit behaviors that can cause distress or aggravation *in others*. Inhibition of behaviors may serve a deceptive purpose, as in sneaky primates who suppress copulation cries so that the alpha male cannot hear and intervene (Dunbar, 1996), and inhibition may serve altruistic behavior, as in a smoker who refrains from smoking in the presence of non-smokers. More importantly however, this kind of inhibition benefits from empathy -- that is, an understanding of how another would perceive and respond to the behavior in question. Moreover, language can play a significant role in enhancing inhibition.

For example, an individual may use a word to calm himself in order to inhibit temper that would bother others, or by using the language of imperatives, he can tell himself (or others) to inhibit behavior that would trouble or hurt other people.

Related to inhibition, Forgas & Ciarrochi (2002) have found in preliminary studies that some successful mood management methods employ information processing strategies. This concerns "affect infusion" (Forgas, 2001) where emotionally charged or affected information intermingles with judgment, thus coloring the outcome of the judgment. Hence, these factors interact complexly: affect infusion, behavior, and the cognitive and linguistically based appraisal of affect and behavior. Moreover, studies imply that cognitive information-processing abilities, which include linguistic processes, enable us to manage and therefore inhibit emotion.

Different cognitive information-processing strategies may not only mediate affect infusion, but could also function as an effective and self-correcting affect management system, producing affect congruence or incongruence in our thoughts and associations to regulate and "tune" our everyday mood states. (Forgas, 2001: 276)

Along these lines even clearer evidence suggests that we can manage our thoughts through self-talk and thus influence our behavior. For example, cognitive behavior therapy (CBT), (Beck and Freeman, 1990) theorizes that in a normal brain, thoughts primarily cause feelings and behaviors; therefore, changing our thoughts enables us to change our feelings and behaviors. Specifically, CBT teaches subjects to see the correlation between their thoughts, emotions, and actions; it requires them to become aware of affect infused or misguided thoughts, and it asks them to perform reality checks on such thoughts. CBT then requires subjects to replace these thoughts with more balanced cognitions and make plans to develop new ways of thinking. Meta-analysis of CBT shows that it has a significantly positive effect on anger management (Beck and Fernandez, 1998). Additionally, CBT has produced well-documented success in treating depression, panic disorder, and agoraphobia, and it has produced encouraging results in treating post-traumatic stress disorder, generalized anxiety disorder, and social phobia. Moreover, results are still pending on array of other disorders, including criminal recidivism (Enright,

1997). Hence, CBT has become the treatment of choice (often in combination with other methods) for numerous mental-health disorders, and it finds increasing support in clinical research studies.

This discussion also raises the question about the difference between language and thought. What aspects of morality can we maintain by thought alone, and what aspects require language? Are there private moralities that we maintain by thought alone? Future research in cognitive science and philosophy may reveal firm insights that show important differences between language and thought that may have some bearing on the questions this thesis answers. Nevertheless, this thesis generally assumes that when we have moral thoughts, we will usually use language to think these thoughts as we internally express them to ourselves. And though we may think morally in private and non-linguistic ways -- through empathic emotions or intuitions, these "feelings" will often also pop up in our heads in empathic words like "Poor guy" or "What a pity that is!" In this way, our private morality may express itself in non-linguistic feelings of empathy, and we may show these feelings to others without using words. Perhaps we can also elicit an empathic response in others by showing them non-verbal empathy as well, but we may also find it hard to feel and act on this empathy without language-expressed thoughts of empathy occurring in our heads. Moreover, though we can feel, think, and show non-linguistic and moral empathy, language will also enable us to express, extend, control, or limit the focus of this empathy.

In addition, other research shows that language appears to enable and sustain the complex interactions between inhibition, altruism, morality, and especially empathy. For example, Deacon and Goodenough (2003) claim symbolic language makes human self-awareness possible.

When we say that we are aware of our thoughts and ideas and plans and memories, we do this using symbolic constructions. It may be possible to have a thought without linguistic representation, but we know that we have had one only when it is self-represented in symbolic form. (Goodenough and Deacon, 2003: 812)

There are ambiguities and problems with this claim, for example concerning the self-awareness of animals or pre-linguistic infants. However, Deacon and

Goodenough maintain that "once language is in place, there emerges not only symbolic reference but also symbolic self-reference in the sense that we humans experience" (2003: 813). Moreover, self-reference combined with empathy extends to various forms of other-reference.

As our vocabularies mature and our ability to manipulate concepts grows more complex, we become able to articulate empathic connection with such abstractions as "life itself" or "the planet Earth." Moreover, we can engage in conceptual blending and configure empathy in radically new ways, as "Love thine enemy." (Goodenough and Deacon, 2003: 814)

Hence, for Deacon and Goodenough, morality comes from our natural prosocial orientation combined with our ability to represent it symbolically. That is, language makes self-awareness possible; language supports prosocial dispositions, including empathy, and language thus expands our other-awareness. Additionally, they claim that once we have made symbolic access to empathy, we can blend it with our predisposition towards reciprocal altruism, or strategic reciprocity. Therefore, "Strategic reciprocity fused with humaneness [empathy] emerges as a sense of justice" (Goodenough and Deacon, 2003: 816). Thus, linguistic ability helps us combine empathy and reciprocity, and it can enhance our basic social and emotional sense of fairness, (Brosnan and De Waal, 2003) for if we empathize and reciprocate with others, these two actions tend to require us to be fair as well. In a word, language powerfully determines our sense of self, our sense of others, and our ability to empathize with them; moreover, language helps us inhibit egoistic behavior and thus engage in altruistic or moral actions.

Now we can focus all these factors regarding self-awareness, empathy, inhibition, language, and morality on the more specific case of cognitive behavioral therapy. For example, CBT relies on subjects keeping written diaries of affective situations, thoughts, and ideas. CBT's success implies that it also provides a means for people to take greater control of their behavior -- including altruistic behavior. To elaborate, since emotions such as anger and empathy clearly impinge on altruism and morality, CBT shows how language helps people manage emotions and behaviors towards empathic responses that positively affect others. For example, CBT can help an individual diffuse his anger by analyzing frustrating events in a

journal and sorting out solutions through writing. In this way, he can avoid the power of adrenaline to hijack the brain and behavior (affect infusion) by employing a cognitive process that works like a sophisticated form of "counting to ten."

In addition to how empathy and reciprocity combine to ratchet up altruism and morality, empathy also forms an important component in human moral codes. Though empathy expresses itself in lower animals as distress at the distress of others and in higher primates as comforting behavior, for humans it refers to the ability to identify with the feelings of another person -- *to feel into another*. This empathic sense, with its biological and cognitive roots, has been extended into a moral code, generally described as the "ethic of reciprocity" or Golden Rule (GR), which is often summarized as "do as you would be done to." Since basic empathy appears fundamental to the constitution of many animals, it is significant that nearly all cultures linguistically formulate empathy in some form of the GR. There are negative or positive versions of the Golden Rule, such as: "don't do as you wouldn't be done to" or "do as you would be done to" in Buddhism, Christianity, Confucianism, Islam, Hinduism, Judaism, Taoism, Sikhism, and Zoroastrianism, and they have been expressed by the likes of Aristotle, Plato, and Socrates. Incidentally, Chapter 3 provides further discussion on the Golden Rule.

Though we do not need to unearth and analyze all the original sources of the GR, it also may not even be possible to do so.

No one has yet dealt satisfactorily with the question of why this precept should have appeared in the codes and outlooks of so many diverse peoples and sages. The golden rule, in one version or another, has had a prominent place in all the major religions and most minor ones; it has been enunciated by pagan philosophers both before and after Christ and by Sophists (Isocrates) and anti-Sophists (Aristotle). There are no detectable historical traces, which could explain this, and the historical diffusion theory is worthless as an explanation here. The nearly universal acceptance of the golden rule and its promulgation by persons of considerable intelligence, though otherwise of divergent outlooks would therefore seem to provide some evidence for the claim that it is a fundamental ethical truth. (Edwards, 1967: 366)

This is vital to our discussion for two reasons. First, the GR *describes* linguistically a bio-psychological phenomenon; that is, it linguistically expresses a moral universal that coheres with the biological substrate of visceral empathy. Second,

language *extends* the bio-psychological substrates of the Golden Rule into explicit cognitive empathy and behavior. Preston and de Waal maintain:

Higher cognitive faculties can also augment helping behaviors through explicit teaching. In human children, direct instruction, reasoning, discipline, and reinforcement of helping are necessary for prosocial competence... These explicit factors can either push a below-threshold state of empathy into an act of helping or subvert empathy altogether when it is required by social rules, but not naturally evoked. (Preston et al., 2002: 20)

Caregivers who teach the GR to a child could *potentially* enhance her prosocial empathy if they socialize her in positive and healthy ways, but it would be counter-productive if they simultaneously teach her the GR but dysfunctionally socialize her. Moreover, Preston and de Waal state explicit instruction (and thus linguistic factors) could enable a person to ratchet up "below threshold" empathy through self-talk or an interlocutor's encouragement. Moreover, this instruction could help increase altruistic behavior. More importantly, the GR instruction may even help people *act empathically* whether they feel empathy or not. In this case, if the non-empathizer still acted altruistically, you would have a "pretend empathy" with the real benefits of empathy. We can express this kind of altruism in phrases like "Don't feel yourself into a new way of acting; act yourself into a new way of feeling," or in Shakespeare's Hamlet "Assume a virtue if you have it not." Thus, to conclude this section:

The complex social world of primates requires the central nervous system to perceive the facial expressions, body postures, gestures, and voices of conspecifics accurately and quickly in order to generate a response. (Preston et al., 2002: 42)

In sum, this section echoes previous sections, but specifically focuses on empathy and language. Language provides a significant ratcheting mechanism, increasing the complexity of empathy in the cognitive realm, enhancing it behaviorally and extending it in numerous ways, in planning empathy, in displacement empathy, and in eliciting empathy from others. Language enables humans to abstractly articulate natural substrate-empathy in phrases like "do as you would be done to." That this phrase has found clear, independent, and ubiquitous expression in diverse cultures correlates well with the fact that it appears intrinsic to a well socialized and

soundly nurtured and empathic mammalian primate nervous system. Incidentally, the Golden Rule assumes this healthy ontology, for unhealthy physical and sociological nurture would skew one's ability to "do as he would be done to." In addition, because we can articulate empathic altruism in the GR, we can explicitly teach it, and teaching empathy means we may produce empathic-like behavior -- whether it is emotionally present or not. That is, even though it is difficult, one can act rationally on a GR principle before or without feeling genuine empathy.

These claims do not deny that humans are capable of atrocities -- or even using empathy against people: "I know this hurts you, and that's why I'm doing it." Ironically, however, leaders often persuade people to commit atrocities for some moral (but false) reason, and though these Orwellian nightmares exist where people use language and lies to diminish empathy, the opposite is also true: we may enlighten and extend empathy, altruism, and morality through language. Thus, "morality without empathy is oxymoronic (Goodenough and Deacon, 2003: 815), and empathy without language makes morality myopic by limiting it from traversing into new categories that transcend biological selfishness and strategic (but still selfish) reciprocity.

2.11 Cooperation, Fairness, Trust, Play, and Language

In addition to empathy, Bekoff (2001; 2002) gives a different description of altruistic behavior that he categorizes as cooperation, fairness, and trust and which he thinks undergirds the evolutionary and biological basis of morality. Bekoff's research concerns the *fair* behavior in the rough and tumble social play of infant dogs, wolves, and coyotes that...

... has its own special rules of engagement, allowing participants to reinterpret acts that might otherwise seem aggressive or sexual. The fact that play rarely escalates into all-out fighting is a strong indication that animals do indeed abide by the rules and that they expect others to do likewise. (Bekoff, 2002: 36)

Canids utilize three special signals that help prevent conspecifics from misinterpreting playful actions. (1) A *bow*, crouching on forelimbs with upright posterior, can initiate play, and (2) stronger playmates show self-imposed

handicaps that limit their force against a weaker playmate, and (3) *role reversal* commonly occurs when a dominant animal allows a subordinate to dominate play. Animals will ostracize those who cheat on this system of rules; hence, cheaters lose playmates. Moreover, animals that fail to play well also bond less effectively to their group, and they are more likely to live outside its protection, and these outsiders tend to die sooner than insiders, implying a correlation between individual reproductive fitness and fair play. Hence, cheaters prosper less. Bekoff thus generalizes that many animals commonly possess a sense of fairness, for social play would not be possible without it.

These issues related to play cohere with previous discussions regarding the importance of healthy socialization for human morality. Though it is beyond the scope of this thesis to detail the relationship between play, socialization, language, and morality, the additional significance of play seems obvious. Bekoff's work analogously supports this, for play among canids forms another important basis for socialization, this time with peers as well as caregivers. Bekoff elaborates:

During early development there is a small time window when individuals can play without being responsible for their own well-being. This time period is generally referred to as the 'socialization period' for this is when species-typical social skills are learned most rapidly. It is important for individuals to engage in at least *some* play. All individuals need to play and there is a premium for playing fairly if one is to be able to play at all. (Bekoff, 2001: 85)

Social play exerts an important influence on human physical, cognitive, and emotional development (Blackmore, 2000) and appears to allow players to train for unexpected events (Spinka et al., 2001). However, the main focus of this brief section concerns play, language, socialization, altruism, and morality. Since play augments socialization, how does human language increase the complexity of play, and how might *language with play* enhance human altruism and morality?

Rules obviously guide play, and the most basic rule, "if you cheat, I won't play with you" does not require language because we can simply stop playing with cheaters. Nevertheless, language can enhance play, because fair play socializes us positively, and because we use language to augment our sense of fair play. First, language enables humans to control play. For example, caregivers often guide the

play of children with language, saying things such as "No!" "Don't touch that!" "Be careful." "Good job." These phrases parallel the instinctive bows, handicaps, and role reversals that regulate the play of canids; however, they also extend the control of play because they allow a person external to the play to control it more efficiently. That is, when the players *behave*, words alone can regulate the play, but if players *misbehave*, caregivers can intervene non-verbally and verbally.

Moreover, the play controller can also control the play of many individuals, or groom the play, (Dunbar, 1996) with greater efficiency than she could without language. In fact, it follows that we could not manage the complexity of play in large groups without language. Second, besides controlling play, language also enhances the planning aspects of play. That is, with language we can *vocally invite* people to play; we can *verbally choose* playmates; we can *orally dictate* modes and places of play; we can *articulately explain* the rules of play, and most importantly all these functions of play come into being or experience greater complexity through language.

Third, next to these factors stands displacement, which as a design feature of language, enables us to displace play from the here and now. Hence, we can imagine, rehearse, and train for play mentally through image training or visualization (without actually physically playing), and language enhances these imaginary aspects of play. For example, children in the middle of imaginary play often narrate their play, or coaches explain and plan the intricate plays of American football, and players execute those plays based on the verbal explanations. Moreover, linguistic displacement also allows for play to take place at a distance, in situations where players cannot see each other, such as hide and seek, correspondence chess, or Internet gaming. Human socialization takes many years, and language continues to develop in normal children through approximately the eleventh year. Thus, play guided by language influences the cooperative, normative, and moral socialization of players.

In sum, fair playing individual canids that do not have language still experience healthy socialization and group protection. For humans, language enhances and increases the complexity of play socialization, thus augmenting the ways that play

protects individuals as they integrate in groups, protecting fair players from the dangers of cheating and banishment. For example, a footballer that plays by the rules can enjoy the protection of his team and league, thus allowing him to stay on the team, play in games, and make a salary. The player who cheats in games and gets red-carded often runs the risk of losing his position, his status, and his pay. This is linguistically guided play, and like gossip it works like grooming (Dunbar, 1996) because it oils the gears of group relations, creates an environment where strategic reciprocity functions more efficiently, and where individuals learn to cooperate altruistically in a group.

2.12 Natural Evil and its Relation to Morality and Altruism

To conclude this chapter on the relation between language and altruistic and moral ontogeny, we need to look at the struggle for survival. This concerns not only how we produce and nurture altruistic and moral behavior but also how we make moral judgments in a world "red in tooth and claw." Despite the substrate altruism and fair play that exists in animals and despite how human language helps ratchet these natural substrates to higher levels, we still see much waste, ruin, selfishness, hatred, violence, and slaughter in nature. Darwin considered that these forms of natural evil justified the theory of natural selection, and modern scholars continue to point to the existence of natural evil to support the theory today. For example, Trivers (1985) cites examples of gang murder and orphan attack in elephant seals, spiteful harassment of injured dominant males by subordinate male mountain sheep, and infanticide of young Langur monkeys by adult males. For Darwin, he reflected on the data he gathered, contrasting his theory with the prevalent creation-perfection paradigm of his time, claiming that only natural selection could explain the data he observed. The following excerpt from the *Origin* exemplifies Darwin's approach for dealing with the problem of natural evil and waste.

As natural selection acts by competition, it adapts the inhabitants of each country only in relation to the *degree of perfection* of their associates; so that we need feel no surprise at the inhabitants of any one country, although on the ordinary view supposed to have been specially created and adapted for that country, being beaten and supplanted by the naturalized productions from another land. Nor ought we to marvel if all the contrivances in nature be not...

absolutely perfect; and if some of them be abhorrent to our ideas of fitness. We need not marvel at the sting of the bee causing the bee's own death; at drones being produced in such vast numbers for one single act, and being then slaughtered by their sterile sisters; at the astonishing waste of pollen by our fir-trees; at the instinctive hatred of the queen bee for her own fertile daughters; at ichneumonidae feeding within the live bodies of caterpillars; and at other such cases. The wonder indeed is, on the theory of natural selection, that more cases of the *want of absolute perfection* have not been observed. (Darwin, 1859: 472, italics mine)

Natural selection follows from the struggle for existence; and this from a rapid rate of increase. It is impossible not to regret bitterly... the rate at which man tends to increase; for this leads in barbarous tribes to infanticide and many other *evils*, and in civilized nations to abject poverty, celibacy, and to the late marriages of the prudent. But as man suffers from the same physical *evils* as the lower animals, he has no right to expect an immunity from the *evils* consequent on the struggle for existence. (Darwin, 1882: 142, italics mine)

Darwin refers to the lack of perfection and the presence of evil in nature, and this relates to our discussion on morality and language because Darwin rhetorically employs moral -- as well as aesthetic, teleological, and theological notions. That is, if we say that something in nature is immoral or evil, then we need a standard by which to make this judgment. Moreover, to do aesthetic assessment, we must first contemplate values pertaining to the rules of beauty, and to make moral judgments, we must conceive of moral notions in the first place. To assess something's teleological status, we must cast a criterion that clarifies what purpose is or is not, and in order for us to theologize -- as Darwin does negatively throughout the *Origin*, we must first envisage positive or negative theological concepts. That is, when we say something does not mesh with a putative divine order, we must conceptualize that order in order to make such judgments (Nelson, 1996; Hunter, 2001).

These theoretical issues raise numerous problems for this discussion because it becomes difficult to ground them empirically, for they require *a priori notions*. Thus, we must acknowledge our presuppositions. For example, when we characterize the empathic and playful behavior of animals as moral-like and their spiteful, hateful, and violent behavior as immoral-like, what is the basis of such judgments? Moreover, the Golden Rule, removed from the context of a healthy nurture or a presupposed ethical system that expects fairness, can become twisted

and negative: "Hurt others as you would like to be hurt by them," or "He who has the gold makes the rules." Hence, these problems lead to difficult ethical questions that I will discuss in Chapter Four, which concerns language and morality beyond group selection and also in relation to Bhaskar's critical realism (1997b).

Nevertheless, briefly regarding moral judgment, Midgley (Midgley, 1991: 3) states: "Asking where ethics comes from is not like asking the same question about meteorites. It is asking *why* we should now obey its rules" (italics mine). This is a teleological question that we cannot simply answer by saying "because it promotes survival," for as we will note in later chapters, some of the most highly esteemed moral behavior comes at a loss of fitness. In this light, we cannot consider as a posteriori the teleological and theological themes in Darwin's Origin because he presupposes non-teleology, and thus he interprets his data in light of these presuppositions. Therefore, with a non-teleological approach, how can we call behaviors or events "evil" without subtly inferring this conclusion from a sense of good and therefore purpose? Strict adherence to a non-teleological presupposition only allows us to say that these behaviors exist -- not whether they are good or bad, and if we say they are good for survival, then we must still explain behaviors that we regard as highest virtue that still do not serve fitness. In short, the vital point here concerns how *a priori* notions will influence any serious attempt to find a discernable foundation in the real world for ethical judgments -- something that I also briefly address in Chapter Four.

At this point, we will set the background for discussion on moral judgments by looking at how "natural social dispositions" relate to morality.

They do not constitute it, but they surely do contribute something essential to making it possible. Do they perhaps supply, as it were, the raw material of the moral life -- the general motivations which lead towards it and give it its rough direction -- while still needing the work of intelligence, and especially *speech*, to organize it, to contribute to its form? (Midgley, 1991: 9, italics mine)

Social instincts conflict with other natural instincts, and from this conflict, a Golden Rule might emerge where we use language to organize our thinking about pro-sociality.

Along these lines, the following summarizes Darwin's discussion of this conflict metaphor in Chapter Four of *The Descent of Man*. First, sociality exists, founded on instinctive empathy and cooperative fair play, and focused on kin. Second, an intelligent creature could remorsefully remember past actions, especially when he yielded to some asocial instinct in conflict with his social instincts. Third, language would enable social animals to express group wishes to guide behavior, and language would help social animals verbally approve or disapprove of behavior. Thus, members would depend on empathy to understand this approval or disapproval as they empathize with the social feelings and judgments of others. Moreover, Midgley (1991) elaborates that narrower social instincts directed at kin for basic socialization would become systematically widened as individuals recognize their expanding social responsibilities that come with group membership. Hence, morality emerges and expands as these narrower and wider issues conflict with each other, and as the group prioritizes its rules, values, and morals...

... not just because they make society smoother, nor even just to make it possible at all, but also more deeply, to avoid lapsing individually into states of helpless, conflict-torn confusion. (Midgley, 1991: 11)

Therefore, when we recall the natural basis for morality in empathy and play, we also see that societies display numerous universals -- including morals and norms guiding group behavior, with language playing a central role in maintaining those norms. We see one of those universals in the "sympathy for others that is generalized by the Golden Rule," (Midgley, 1991: 12), and this raises the question, as Midgley does, whether a society is even possible without the GR. The benefits of individual socialization and group membership would not exist if the wider and more empathic instincts did not at least partly overrule selfish gene instincts. If so, the Golden Rule exists universally because it helps cultural groups not only survive fitness-wise, but also thrive emotionally and socially.

Hence, an immoralist in this context becomes one maladapted to group life whose behavior shows disregard for empathy and group cooperation and deserves to be labeled as wrong. Maintaining group relations concerns resolving conflicts between narrower and wider interests and being able to create shared solutions. With the constraints of group size, this would only find feasible solution with language.

Thus, "as much as language, [morality] seems to be something that could only occur in naturally social beings" (Midgley, 1991: 12), social beings who not only want to survive, but who also employ moral systems to achieve well-being, happiness, and peaceful coexistence. In short, this discussion of morality and language leads us to the importance of group cohesion and group selection, the focus of the next chapter.

CHAPTER 3

LANGUAGE, MORALITY, AND GROUP SELECTION

The previous chapter developed themes regarding the necessity of language to develop and maintain altruism and morality for human individuals and groups. Up to now, we have focused briefly on why altruism and morality present problems to evolutionary theory and how these problems have stimulated evolutionists to develop ingenious theories that successfully explain altruism to a degree. Nevertheless, this chapter argues that though mainstream theories impressively explicate many forms of altruistic behavior, some kinds of altruism may require additional explanatory mechanisms. For example, if SGT fails to elucidate forms of altruism like sacrificial philanthropy, then we need a different searchlight to explore such unlit contours of human behavior. Group or multi-level selection theories may help us understand behaviors that SGT can only explain incompletely. Moreover, evidence may point to group selection (GS) operating on human groups who possess language. Language could unify a group to the degree that individuals *virtually* act together as a single organism, enabling them to sacrifice fitness for the group, and this could cause the group to become subject to GS pressures. Thus, a main question is: do moral codes exist in groups that regulate altruistic behavior in a sufficient number of individuals to cause the group to out-compete an opposing group of selfish individuals?

3.1 Altruism and Morality: A Central Problem for Evolutionary Theory

The history of the problem of altruism in relation to natural selection is fascinating, and the following section, relying on Schloss (1998; 2002), demonstrates the

appeal and significance of this topic. It is not trifling that a significant train of reasoning in the *Origin* and virtually all major subsequent Darwinian discourse treat the problem of imperfection, waste, selfishness, and evil in nature as confirming evidence for natural selection. Moreover, these issues relate to how Darwin empirically refuted the "theory of creation" prominent at his time, especially as promoted by William Paley. For example, Darwin used terms "creator" and "theory of creation" over 60 times in the *Origin* and *Descent of Man*, to contrast the then prevailing theory with natural selection.

This is significant because the moralizing of his era was religiously based, and though he acknowledged the moral role of religion in society, Darwin contended for a natural origin and grounding for morality. This generated controversy, and still does, but besides the hullabaloo (and inherent interest) of the debate, the religious and scientific elements of this discussion on morality are extremely important. Pinker (2002) and Wilson (2002) demonstrate this by prominently attending to these issues as they articulate their different evolutionary views of human nature and morality. Moreover, because natural selection explains away evil and waste in nature, it sets up a contrary problem in light of the selfish solution it provides. That is, with waste, selfishness, and evil as by-products of natural selection, another opposite problem now stares us in the face. As D. Oates states:

For followers of Darwin, the familiar theological 'problem of evil' was turned inside out: evil could henceforth be assumed, and the existential paradox, which demanded explanation, became, in fact, *the problem of goodness*. (Oates, 1988: 444, emphasis in the original)

3.2 Solutions for the Problem of Altruism?

This problem and the efforts made to solve it have given us two important advances in evolutionary theory: kin altruism and reciprocal altruism. To Review, Hamilton's (1964) kin selection became the first major advance towards untangling the "problem of goodness" from the perspective of natural selection. Hamilton demonstrated that altruistic and sacrificial behavior directed toward offspring ensures an organism's fitness by maximizing its genetic contributions in following generations. The theory predicts that organisms will altruistically invest in

conspecifics' welfare proportionally to their genetic relatedness. Moreover, kin selected altruism successfully explains previously non-understood sacrificial behaviors in numerous species and has garnered wide empirical verification (Trivers, 1985). Hence, kin selection makes clear sense of behaviors that oppose individual reproductive fitness, but favor the fitness of the altruist's genes in kin. Trivers (1985) claims kin selection is the most important advance in evolutionary theory since the fusion of Darwin and Mendel, and Hamilton's work is thus considered a major synthesis in evolutionary theory.

Next to kin selected altruism, reciprocal altruism (Trivers, 1971) became the next major advance to synthesize natural selection and socially sacrificial behavior. Reciprocal altruism is the flip side of "tit for tat" reciprocal revenge, which says, "I'll damage your fitness as much as you damage mine." Hence, in reciprocal altruism, an "eye-for-an-eye" positively becomes, "If you give me an eye's worth of fitness benefit, I'll return the favor." We can say that reciprocal altruism exists in other social organisms, for example, in symbionts -- plants or animals that live in close and mutually beneficial association with different species. However, for humankind, we can empirically verify reciprocal altruism simply because it exists as a cultural universal (Brown, 1991). In fact, the universality of reciprocity in human cultures caused Becker (1956) to refer to humans as *Homo reciprocus*.

Related to reciprocity, it would present an interesting quandary if a form of altruism were actually shown to exist for the exclusive benefit of another species.

Natural selection cannot possibly produce any modification in a species exclusively for the good of another species; though throughout nature one species incessantly takes advantage of, and profits by the structures of others. But natural selection can and does often produce structures for the direct injury of other animals, as we see in the fang of the adder, and in the ovipositor of the ichneumon, by which its eggs are deposited in the living bodies of other insects. If it could be proved that any part of the structure of any one species had been formed for the exclusive good of another species, it would annihilate my theory, for such could not have been produced through natural selection. (Darwin, 1859: 200-201)

Clearly, reciprocal altruism requires give and take, so it cannot exist for the exclusive benefit of another individual or species -- but for mutual benefit.

Nevertheless, Darwin's concern regarding the goodness of traits in one species exclusively benefiting another species merits further comment. First, it does not necessarily follow that such traits would annihilate his theory, though they might limit its scope regarding some kinds of sociality. Besides this, it appears that most - if not all -- inter-species symbiosis is reciprocal anyway. Second, though Darwin showed concern about inter-species unconditional altruism not fitting with his theory, and though he mentions group selection, he does not mention the narrower issue of non-reciprocal altruism benefiting members of the same species. Could natural selection produce traits in individuals that exist for the exclusive benefit of another individual of the same species? This is extremely unlikely by the theory and the evidence, but an even narrower problem remains.

The most significant problem facing reciprocal altruism is that there exist numerous observable behaviors in plants and animals, and especially in humans, that are inexplicable on the basis of either kin selection or reciprocal altruism. Human behavior regularly exhibits its noncompensatory sacrifice for nonkin. Explaining such phenomena constitutes one of the most controversial aspects of current evolutionary theory (Schloss, 1998: 242-243)

Simply put, humans and other species can behave in altruistic ways that do not give them a reproductive or compensatory benefit; therefore, kin selection and reciprocal altruism cannot explain all forms of altruistic behavior. Of course, we will need to test this claim against behaviors that seem genuinely altruistic, as in the following list from Schloss (1998).

- Sacrificial philanthropy.
- Blood and organ donation.
- Heroic and one-time efforts to save others across kindred lines.
- Acts of wartime or athletic bravery incurring risk. (Some athletic behavior may incur a fitness risk, and though it could be done to promote the actor's status, it could also be done for the team, i.e., the group).
- Exclusive or monogamous homosexuality. (Since this behavior produces no genetic fitness benefit, it could possibly in some cases be a form of non-reciprocal altruism).
- Suicide. (For example, some forms of Japanese ritual suicide are done for the honor of someone else, for one's master, or group, so theoretically, this could be a form of sacrificial altruism.)

We might explain some of these behaviors from a selfish gene perspective: sacrificial altruism might confer status on the actor, and this would grant him

greater access to resources that would benefit fitness. For example, Zahavi and Zahavi (1999) claim that a peacock may risk dangerous tail displays to show females that even with such a handicap that it is a strong genetic survivor with the best genotypes. However, for humans we may not know their motives for sacrificial altruism, and some people may secretly behave altruistically, and this would remove the possibility of gaining status or resources from their behavior. Because of these kinds of factors, we should seek additional explanations also. Thus, for the purposes of this chapter on group selection (GS), we will evaluate a few of the behaviors listed above to see if GS can explain them. This is important because selfish geneticists tend to interpret such behaviors as anomalies; however, before looking at GS and the reasons for invoking it, we must consider how these genic selectionists have explained away atypical altruism.

3.3 Explaining Real Altruism as Anomalous

Darwin's idea of sexual selection stands as one approach that shows how apparently counter-reproductive behavior might ultimately add to reproductive success and provide a way for anomalous traits to serve fitness. For example, the peacock's tail makes the bird vulnerable to predation but also potentially increases its chances for mating attractiveness and success -- through this form of 'sacrificial' behavior. Moreover, altruistic behaviors that enhance social status may serve as another benefit of sacrificial altruism that would increase individual reproductive fitness (Dessalles, 1998). In short, traits that endanger fitness might on the balance benefit sexual attractiveness and selection, and increased social status can attract mates resulting in sexual selection, and higher status could also benefit reciprocity or contribute to dominance and increased resources.

However, there are several difficulties in using sexual selection and the role of status to dismiss real altruism. First, a sacrificial altruist may benefit reproductively from his behavior, but this is not guaranteed, especially if he dies in the process. Second, Schloss (1998) remarks that the genetic basis for these behaviors is at best unclear and at worst tenuous; moreover, empirical studies do not exist correlating these sacrificial behaviors with social status, nor with fitness in human populations.

Even if we grant that SGT can explain these altruistic behaviors, this still does not show that SGT explains them most accurately or parsimoniously. In addition, we cannot hastily resist the criticisms of Lewontin (2000), Gould (2001), Rose (1997) and other Darwinists who have raised concerns about employing selectionism as a theory of every kind of behavior. In short, the rejoinders here do not refute SGT explanations of apparent altruism, but they point to problems that call for open-mindedness towards other relevant theories.

Nevertheless, besides status and sexual selection, there are still other ways to explain counter-reproductive altruism that we need to address before we can justify additional explanatory tools. These additional SGT explanations include pleiotropic interactions and genetic lag. First, regarding pleiotropy, where one gene determines two or more apparently unrelated characteristics in an organism, the "maladaptive characteristic" of atypical altruism may emanate from the adaptive traits of a related genotype or phenotype. For example, a pleiotropic interaction may cause us to respond empathically and altruistically when we identify and recognize an infant's face, and this could be a maladaptive response (from the gene's eye view) if we direct this altruism to non-reciprocal non-kin. Moreover, SGT may interpret adoption of children as atypical altruism that does not increase reproductive success but that still exemplifies a maladaptive expression of parental care. In the adaptive form, parents behave altruistically to the genetically correct natural offspring, but in the putative maladaptive form parents treat adopted non-kin as if they were offspring.

Second, genetic lag refers to the idea that particular behaviors conferred a fitness benefit in the environment of evolutionary adaptation (EEA), but not in current environmental conditions. An evolutionary psychologist might explain adoption of children and bravery in war as behavior best adapted to tribal life based in kinship and reciprocity. However, in current societies evolutionary psychology would construe this conduct as maladaptive because warfare and adoption in large societies do not clearly serve kinship ties, nor do they deal in direct reciprocity. Hence, because cultural change outpaces evolutionary change, our adaptations fill a mixed bag of behaviors -- some fit to the current scene, and others

anachronistically out of sync with life in the modern world. Nevertheless, though such EEA explanations for altruism cohere theoretically, they clearly lack strong empirical support.

One other explanation of maladaptive counter-reproductive altruism that meshes with selfish gene theory concerns "manipulation and the extended phenotype" (Dawkins, 1982; Schloss, 1998). In this scenario, an egoist manipulates altruistic behavior from an altruist in order to serve his selfish ends (Trivers, 1985). Nest parasitism is a prime example, where cowbirds manipulate the parental instincts of warblers, getting them to invest in cowbird offspring and causing them to sacrifice their reproductive success for another species. For humans, the SGT perspective sees adoption and blood and organ donation as counter-reproductive behavior where the reproductively beneficial phenotypes of the altruist are manipulated to benefit non-related and non-reciprocal conspecifics. Thus, Schloss perceptively sums up how the SGT perspective views human altruism.

In this way every act of apparent kindness becomes ultimately self-interested or involuntarily manipulated. In its favor, this view unmasks any saccharine ideal of human beneficence that denies the manipulation of prosocial instincts. (Schloss, 1998: 245)

Humans can clearly manipulate others to take advantage of altruistic instincts; however, there are problems with using manipulation, genetic lag, pleiotropic interactions, and sexual selection to deconstruct every last vestige of genuine altruism. For example, manipulations potentially introduce a *telos* into the interaction. Natural selection designs behavior for one individual's reproductive success that another organism manipulates for its own benefit. Though manipulation fits with selfish gene theory, it is difficult to show exactly what these behaviors are for, and whether they have been manipulated or not. Moreover, it is hard to know how to test for apparently teleonomic processes where the egoist diverts to himself the reproductive benefit of a behavior primarily designed *for* the altruist (Schloss, 1998). In some cases, for instance, the cowbird clearly manipulates the instincts of the warbler, but we should be very cautious to interpret all cases like this without strong empirical support. Most importantly, as we will see below, some cases of radical altruism are very hard to deconstruct with SGT.

In spite of the problems with the ways some researchers will deconstruct atypical altruism, the key emphasis throughout this discussion remains that some real altruistic behavior appears to transcend these deconstructions. However, before looking at these behaviors and how they cohere with a GS perspective, we need to see how selfish selectionism (SS) deals with these problems. For selfish selectionists, the ultimate units of selection and fitness occur at the genetic level -- where phenotypes exist as the means for genes to interact successfully with the environment. Hence, for example, all phenotypic interactions that provide parental care, including manipulated interactions, become extensions of the selfish gene's survival mechanism (Schloss, 1998). That is, SS sees all cases of parental care, either coming directly from parents, or manipulated from others, as phenotypic interactions that provide for the reproductive fitness of the beneficiary genes.

The whole idea of altruism, and the challenge to Darwinism it supposedly represented, becomes an artifact of an organism-centered rather than gene-centered perspective on fitness. Dawkins argues, moreover, not only that we should focus on genes rather than organisms but also that our understanding of organisms themselves should expand beyond the skin -- to entail all the phenotypic interactions between genes and their environments. (Schloss, 1998: 246)

The rhetorical allure of selfish selection theory takes shape in its ability to dispel (and dispense with) the problem of altruism altogether. Organisms in selfish gene theory become a "fluid conceptual construct" (Schloss, 1998: 247) where genes produce phenotypic altruism for themselves. This occurs even when the phenotype that actually performs the altruism exists in another organism -- but as the extended phenotype of the manipulator's genes. Thus, organisms -- the manipulator, the manipulated, the egoists and the altruists -- all disappear from the altruistic calculus because we focus our interpretive lens only on how the egoist promotes his own fitness. Nevertheless, despite the technical problems that many scholars see in SS (Volland and Volland, 1995; Waller, 1995; de Waal, 1996; Rose, 1997; Sober and Wilson, 1998; Dover, 2001; Balazs, 2004), the key point concerns the astute rhetoric of selfish selection.

That is, selfish selectionism simply defines atypical altruism out of the explanatory picture. Surely, inclusive fitness and reciprocal altruism explain numerous

behaviors in nature that support SS in a broad but still limited sense. However, SS subtly defines the problem of altruism in a way that only allows for selfish gene explanations. Thus, if selfish genes can manipulate the phenotypes of other organisms to the degree that these phenotypes become extensions of the manipulator's genes, then it becomes conceptually impossible for atypical altruism to exist. The gene's eye view gives a helpful perspective, but if we employ it as the *only* perspective, we may fail to accurately see and understand some behavior that exist outside of it. Of course, SS could obtain empirically; however, as theory, it also pre-empts the very possibility of atypical altruism.

Hence, SS deconstructs all real altruistic behavior, calling it manipulation or genetic lag; then SS labels this behavior as apparent altruism. Though some manipulation and genetic lag exist, this deconstruction still results in two serious problems. First, some real altruism may truly exist, and second, definitional gerrymandering -- that theoretically precludes the existence of real altruism -- locks research on altruism into predetermined conclusions. Incidentally and ironically, Dawkins suggests a solution to these problems -- that actually contradicts SS theory -- in his proposal for memes -- and this will be the focus of the last chapter of this thesis.

3.4 Justifying an Appeal to Group Selection

Atypical and real altruistic behaviors that transcend selfishness attest to theoretical and empirical incompleteness of selfish selection theory. These altruistic exceptions to the rule of selfishness, especially in light of the restrictive nature of SS, provide good reason to add to our theoretical toolbox for understanding altruism. Nevertheless, these problems do not resolve the *internal and conceptual* issues pertaining to group selection (GS), and therefore they do not completely justify an appeal to GS. If it were that simple, GS would represent a mainstream solution to the problem of altruism in nature. However, Sober and Wilson (1998) have helped to clarify and resolve these conceptual issues, and their work is the main reason GS is gaining broader acceptance. Thus, in light of Sober and Wilson's

work, and the issues mentioned above, the present section summarizes the rationale for reinstating GS into the study of altruism.

Regarding GS, Lewontin states it "has been regarded as anathema by nearly all evolutionary biologists, although entirely without empirical evidence" (1998: 4). Though Lewontin's claim is polemical, it also expresses the contentiousness of the issue and shows the difficulty of proving a negative "that GS does not exist." Moreover, it reminds us of the self-confirming way SS advocates have defined the problem. Though many would admit the theoretical possibility of GS, scholars generally affirm that, "Most features of organisms -- their development, biochemistry, physiology, and life history -- are the product of individual selection" (Stearns and Hoekstra, 2000: 49).

In addition, Dawkins, (1976: 12) promotes a narrower version of selection in which "the fundamental unit of selection... is not the species, nor the group, nor even, strictly, the individual. It is the gene, the unit of heredity." In his favor, Dawkins argues forcefully that individuals are genetic entities too large and temporary to qualify as significant units of selection, and population groups are too big, unstable, and non-unitary for selection over other groups. Biologically speaking, therefore, there is potent force to the argument that genes produce the best copying, mutating, and adapting mechanism for natural selection. Moreover, because the fecundity, copying-fidelity, and longevity of genes span the total history of life, and because there is a solid scientific knowledge base in genetics, molecular biology, and artificial selection, it makes good sense to reduce the unit of selection to the gene.

For levels of selection, we can contrast individual and genic selection with species and group selection. Species selection refers to the process of sorting between species at different rates of speciation and extinction. Species selection, therefore, happens too seldom to produce adaptations, but rather affects "patterns in the abundance of families, orders, classes, and phyla" (Stearns and Hoekstra, 2000: 49). Conversely, group selection refers to the process of environmental effects, producing varying degrees of reproductive success among groups within species. That is, GS concerns factors that influence traits that confer fitness benefits on

groups, and many biologists consider it at best a weak form of selection for the following reasons. According to Stearns and Hoekstra (2000: 50), individuals reproduce and die more frequently than groups and reproductive performance corresponds more strongly to individual traits than to group traits. Thus, response to individual selection is faster and more precise than response to group selection, and the reproductive success of groups and species arises as a by-product of the individuals in those groups. Hence, the argument goes, exceptions to individual selection are logically possible, but probably infrequent.

Despite these issues, we should note that theories for kin selection, reciprocal altruism, and the selfish gene "were developed to explain the evolution of apparent altruism without resorting to group selection" (Sober and Wilson, 1998: 55).

Developing a theory in order to avoid resorting to another theory is not wrong, but it is inconsistent to say that GS is theoretically possible but empirically unlikely if our theory tends to define GS out of the picture. The definitional aspects of this debate remain significant to the status of group selection because as GS has started to regain acceptance, a main argument against it concerns definitions.

The biologist David Sloan Wilson and the philosopher Elliot Sober recently redefined "group" as a set of mutual reciprocators, providing an alternative language in which to describe Trivers' theory but not an alternative to the theory itself. But no one believes the original idea that selection among groups led to the evolution of unstinting sacrifice. (Pinker, 2002: 259)

Pinker's dismissal of GS reveals instability in evolutionary theories of altruism because it shows how views are split on the subject. Moreover, this lack of accord may also indicate that the question of altruism still stands at the frontiers of such research. Incidentally, even if *one aspect of renewed GS theory* provides alternative language for reciprocal altruism, if the actions of selfish, mutual reciprocators cause a group to out-compete another group, then GS would obtain despite this selfish behavior. Moreover, a definitional debate with SS emerging as the winner tends to make kin selection the sole ratchet of altruism and thus, "altruism [eclipses] adaptation as the central question of sociobiology" (Sober and Wilson, 1994: 158). In any case, we would be wise to follow David Hull's advice (2000). Though referring to a different issue, he suggests we postpone definitional issues;

concentrate on results, which can dialectically refine conceptual conundrums. That is, empirical results -- for or against GS -- will dialectically elucidate definitional concerns, and thus, empirical evidence becomes the most important criterion for developing GS theory and evaluating its status.

Regarding the dismissal of GS, a main problem concerns how SS fails to acknowledge that GS is *natural selection based on the differential fitness of groups*. Thus, biologists may interpret potential examples of GS as individual selection by failing to check for the presence of groups and by failing to check whether a trait in question "evolves by virtue of a fitness advantage within groups or a fitness advantage between groups" (Sober and Wilson, 1998: 52). Sober and Wilson call this the averaging fallacy, the process of averaging individual fitness across groups. For example, instead of seeing a group of individuals united by an altruistic trait, we simply measure the fitness of all altruistic and selfish individuals in a population. In this way, selfish individuals not only win on fitness, but groups defined and unified by altruistic traits disappear because we fail to look for them. However, if we look for altruistic and selfish groups in a population, then the altruistic group (if it exists) will win on fitness against the selfish group, provided it can keep cheating in check. Thus, selfish selectionism pre-empts group selection in two ways: (1) the extended phenotype that turns the altruist's behavior into selfishness, and (2) the averaging fallacy that dissolves the presence of altruistic groupings in the population.

With the averaging fallacy, the "existence of more than one group and fitness differences between the groups have been folded into the definition of individual selection, defining group selection out of existence" (Sober and Wilson, 1998: 32). Moreover, the averaging fallacy not only disallows the possibility of GS on a priori grounds, it also fails to account for opposing forces within groups and between groups. These forces would account for between-group selection, which would favor the evolution of altruism, while within-group selection would reinforce the evolution of selfishness. Hence, the problem of defining natural selection as individual selection fails to account for the possibility of pressure between groups.

Pressure between groups could promote the existence of atypical altruism if the cheater problem and migration between groups can be kept in check.

Though a full defense of group selection goes beyond the scope of this thesis, see Sober and Wilson (1998), we can briefly see some empirical support for GS in female-biased sex ratios that help maximize group fitness. That is, group selection favors female-biased sex ratios that increase the fitness of the group by helping it to produce more offspring (Sober, 2005). This also results in a greater parental investment in females. Individual selection favors what is good for individuals -- equal parental investment in both sons and daughters. Thus, even sex ratios will cause some individuals to thrive, but female biased sex ratios will cause the group to thrive at the expense of some individual fitness.

Moreover, this commonly occurs in hundreds of thousands, perhaps millions of small invertebrate species in ephemeral habitats before they disperse to more permanent habitations (Sober and Wilson, 1998: 41). Regarding female-based sex ratios supporting GS, scholars averse to the idea need to consider at least four important issues. First, averaging fitness across groups melds individual and group selection processes together yielding one summary statistic on fitness. This simply fails to see the existence of groups -- even temporary ones. Second, SS theory couches the issues in such a way that it hinders us from perceiving the existence of conflicting selection pressures. Sober and Wilson clarify:

The biased sex ratios that evolve are *compromises* between the simultaneous and opposite influence of group and individual selection. Group selection rarely, if ever occurs without individual selection occurring as well. (Sober and Wilson, 2000: 194)

Thirdly, Sober and Wilson claim that Hamilton (1967), whose work forms the foundation of SS theory, acknowledged that female based sex ratios empirically support GS, and this admission is not widely known by SS theorists. Fourth, it is beside the point that this scenario for female-based sex ratios may occur only over a short period of time. That would mean that GS operated for a time in that group - not that GS does not exist.

Female-based sex ratios and numerous other proposed examples of GS (see also Dicks, 2000), should encourage researchers to keep an open mind regarding group selection. However, while some leading biologists concur that GS presents a viable selection pressure on human groups (Darwin, 1882; Mayr, 1997; Sober and Wilson, 1998; Wilson et al., 2000; Wilson, 2002), GS will remain controversial for a while because it contradicts current orthodoxy. Furthermore, the fundamental nature of Sober and Wilson's critique may intensify this controversy. According to Lewontin: this critique "is precisely that combination of radical re-examination of a system of explanation, an examination from the roots, with a rigorous technical analysis of both biological and epistemological questions that we all are supposed to engage in (Lewontin, 1998: 6). Nevertheless, despite its controversial nature, we have enough scholarly consensus to justify research into the relationship between group selection, language, altruism, and morality.

Thus, what follows summarizes Sober and Wilson's (1998) argument on how group selection might operate. Difficult environmental conditions would push groups to function as an organism -- where, for example, group hunting and gathering would promote group survival. This in turn would encourage primary and high cost altruistic behaviors, such as sharing food and resources. Groups can encourage these primary behaviors through secondary and low cost behaviors such as rewards and punishments, which express the group's moral system. Punishments would be aimed at selfish cheaters who might take advantage of the altruism of other members. These secondary altruistic behaviors cause and amplify primary altruistic behaviors, and between-group selection would favor codes that benefit the adaptive fitness of the group. In short, although selfish human nature may tend to promote an individual's adaptive fitness within a group, traditional human cultures tend to show a design that hinders selfish selection processes that detract from group fitness. This aspect of group selection is evident in the history of groups displacing other groups. Moreover group fitness emerges from the use of moral codes, which are enabled and extended by language; therefore, moral codes depend on the evolution of language and the adaptive benefit that language bestows on groups.

3.5 The Relation Between Group Selection and Language

Language links up with morality and group selection because human groups use moral language to impose social control. "What is uniquely human is the harnessing of socially shared values" (Sober and Wilson, 2000: 195), and this thesis contends that language provides individuals and groups with mechanisms to create, maintain, and harness a shared set of moral norms. If language helps us create and maintain a moral code that imposes social control that limits the fitness of individuals for the greater fitness of the group, then groups could coalesce and become subject to GS pressures. Thus, this relationship between language, morality, and groups suggests a set of interesting multi-directional interactions.

- How does language influence groupishness?
- How does language influence morality?
- How does morality influence language?
- How does morality influence groupishness?
- How does groupishness influence language?
- How does groupishness influence morality?

These questions show the challenge of teasing out the causal relation between language, morality, and groupishness that we consider below.

3.5.1 *How does language influence groupishness?*

This question concerns how language serves group cohesion. First, a language provides a mutually intelligible system for its speakers, and this shared system gives a group unity that comes from using the same phonological, syntactic, and semantic structures. Roughly the same linguistic software resides in the mental hardware of all the members of a particular sociolinguistic group -- allowing members to efficiently share data. A person who speaks an incomprehensible language belongs to an out-group, and hence, linguistic differences create groupishness and help distinguish one group from another. In the case of mutually intelligible dialects, phonological differences still serve as a softer form of group boundaries.

In addition to ways language links up with groupishness, language also provides access to in-group information. Becoming a part of the in-group not only requires

mastery of a unique code, but it also provides access to the information couched in that code, which insiders use to function within the in-group. Moreover, the ability and privilege to use such information only comes with advanced knowledge of the in-group language. Linguist Donald Larson, (1984) coined the terms B-I Passing and De-Alienation to describe the process that an outsider experiences when adapting to a new cultural group and language. B-I Passing refers to the process of passing little by little into the target group, becoming more and more linguistically and culturally acceptable to the insiders. De-Alienation refers to becoming aware of what insiders find linguistically and socially appropriate or repulsive so that one can act on this knowledge in order to practice B-I Passing. These processes require: (1) decreasing our dependence on outsiders; (2) decreasing our isolation from insiders; (3) increasing our involvement with insiders, and (4) increasing our independence as we seek involvement with insiders. Besides these behaviors, moving towards the inside of a group involves value transformation, which will vary in difficulty and depth depending on how the target group's values differ from ours and how the in-group metes out acceptance.

With these processes in mind, imagine an outsider immersed in a new group, with no knowledge of that group's language and without any information in his native language about the group's ways. His goal of getting on the inside will include linguistic change, but it may often also require behavioral and value changes that will garner him in-group access and acceptance. A sensitive language learner will perceive disapproval for some of his actions through the non-verbal communication of the insiders, but without specific information -- encoded in the language -- he will have trouble knowing what behavior is acceptable. Moreover, the key factor that makes one acceptable is linguistic ability because a non-speaker of the target language is outside the linguistic border of the group. Insiders can use non-verbal behavior or translation to help an outsider get inside the group, (if they see a benefit for letting him in or have come to empathize with him), but he most efficiently can get inside the group when the in-group's language gets inside him.

Hence, B-I Passing and De-Alienation highlight how socio-linguistic groups form bodies or organisms that are difficult for outsiders to enter. In fact, though the

organism analogy may be imperfect, it still clearly parallels the real world in significant ways. For example, when a foreign virus enters a body, physical defense mechanisms protect it from these viral invaders. For human groups, some of these defense mechanisms are intentional (rules to keep foreigners out) and others unintentional (the barrier of language itself), but in both cases language powerfully supports these defense mechanisms.

Group protection mechanisms can be insidious, such as apartheid where separate toilets and transportation systems exist for minorities, or where ushers police churches to insure only people of the 'right' color are in attendance. These systems are based on color, but large groups could not enforce them without language. In contrast, group protection mechanisms can also be nourishing to the in-group defending it from negative outside influences. For example, we certify medical doctors to ensure that only qualified individuals belong to the group of people who legally practice medicine. Incidentally, this code protects both the in-group and the out-group it serves. Language can efficiently establish, enable, and maintain both the insidious and the nourishing forms of group protection to greater levels than a lack of language could allow.

3.6 How does language influence morality?

Besides linguistic influence on boundaries and cohesion, language also influences the morality of groups. This is not the simplistic claim of linguistic relativity that language itself creates the moral values and perceptions of a group. However, it is a claim that language helps develop and maintain morality in individuals and groups, in an essential way -- as discussed in the first chapter. Here are some of the ways groups employ language in relation to morality.

- Groups encode morals and verbally demand adherence to them through language-based statements, like the Ten Commandments.
- Groups transmit this moral code from generation to generation, either in oral or written form, and this transmission is utterly unachievable without language.
- Groups use linguistically encoded promises, threats, rewards and punishments, to enforce moral norms.

- Groups negatively mark moral deviants (i.e., cheaters) through linguistically based gossip or other monitoring strategies. Using language, groups threaten, punish, mark, fine, intern, excommunicate, and in extreme cases command the execution of deviants.
- Groups positively mark moral adherents (i.e., in-group altruists) through linguistically based gossip or other monitoring strategies. Using language, groups reward, bless, praise, honor, and even immortalize those who adhere to group norms.
- Language provides a cost efficient means for classifying behaviors as good or bad, right or wrong, especially when dealing with large numbers of individuals in groups.

These linguistic means bolster morality and have important implications. Explicit "thou shalt" and "thou shalt not" statements allow groups to agree on norms in a way not possible without language. Primatologist Francis de Waal expresses the way that language uniquely allows for social and moral contract building:

There is the consensus building permitted by language. Morality is a group consensus about how conflicts of interest within the community should be resolved. Language may make for a much better coordination between individual standpoints on this, guiding a "democratic" consensus building process, and permitting more clearly for individual differences to be expressed. (de Waal, 2002: personal communication)

Linguistically weak outsiders experience great difficulty when trying to make their interests known in a foreign language, and even native speakers who share the same language can fail to communicate. As selfish selection shows us, behavior that benefits one's genes and offspring is more common than its atypical altruistic counterpart. Selfishness may be the innate norm, yet humans must survive in groups of selfish individuals. For these groups to thrive, they must have a means for efficiently, effectively, and articulately building consensus to manage conflicts of interest that naturally arise in groups of selfish individuals. Physical force can help deal with conflicts of interest, but it disadvantageously incurs high fitness costs to all parties involved. However, language helps groups explicitly agree upon a normative code, a code that can help enforce the will of the group and limit the costs of enforcement. Moreover, while enforcing the will of the group always incurs cost, to attain consensus, linguistic moral assertions will cost less than physical ones.

Moreover, though verbally agreeing on norms may incur little cost, the means of enforcing these norms will incur varying degrees of cost depending on how the group imposes them. Related to this, Sober and Wilson (2000: 195) suggest a scenario for how groups use "cultural norms that impose social controls." They consider the costly act of group members donating ten percent of their food to the whole community, and this act requires a high correlation of donation behavior between group members to maintain it. However, when the group promises to reward altruistic donors and threatens to punish selfish hoarders, then rewards and punishments make it "no longer altruistic to make the donation, but selfish" (Sober and Wilson, 2000: 195). Groups who punish cheaters decrease cheater fitness, and groups who reward donors increase donor fitness. Though social controls may increase the selfishness of the altruistic acts, these acts remain altruistic for how they affect the group. Hence, increasing the average food resources for all group members also increases the fitness of the group enabling it to out-compete other groups, particularly groups who do not employ the ten percent donation rule.

In this scenario group selection comes into play. Though the donation becomes less altruistic because of the linguistically enabled punishment and reward system, those who enforce the donation will risk fitness due to the danger of punishing a cheater. Therefore, as the donation becomes selfish, enforcement becomes altruistic, and though the group could reward enforcers for their altruism (thus adding a selfishness into these altruistic acts), enforcing is still altruistic. Hence, despite the individual fitness risks for those who enforce punishments, the donation system appears to translate into group fitness. In addition, we can more easily produce the enforcing behavior in the group (to evolve the donation behavior) than trying to develop altruism without rewards and punishments. Thus, enforcing linguistically based norms becomes a *relatively* easy altruistic act that evolves a *comparably* more difficult form of altruism by adding selfishness to it through reward. This kind of altruism can translate into group fitness and therefore group selection.

Not only do groups encode, reward, punish, and enforce moral norms, they also employ language to mark deviants and cheaters through gossip. Wilson and others

(2000) claim that group-serving gossip can function as a social control mechanism that serves as a group level adaptation. Group-serving gossip activates social control by helping groups negotiate their values -- "so-and-so is doing X and thus not living up to our standards," and this verbal act punishes or banishes those who violate group norms, thus defining the boundaries of group membership. Hence, if a behavior makes a person unacceptable to the group (including self-serving gossip), group-serving gossip verbally categorizes him as an outsider, thus placing him outside the boundaries and blessings of the in-group. Gossip helps groups monitor and exclude cheaters and reserve the benefits of group membership for cooperators. It may also serve verbal group grooming (Dunbar, 1996). Groups who limit cheating and increase altruism through verbal gossip may increase group fitness in comparison to other groups who do not maintain altruism in this way.

Regarding group selection, we need to keep in mind the difference between group-serving and self-serving behavior. Group-serving behavior brings a payoff to the altruist that cannot be measured by comparing fitness between individuals within or between groups but rather by comparing how well a group functions in relation to other groups. (This avoids the averaging fallacy). The key point here concerns how gossip serves as a social control mechanism that promotes group-serving behavior and suppresses self-serving behavior. For this to happen, Wilson et al (2000) assert that gossip must meet these four conditions. (1) It must help groups detect self-serving, non-group-serving behavior; (2) it must impose a high fitness cost on those punished; (3) it must permit punishers to perform their duties at low cost, and (4) gossip must be context sensitive, so groups can mark self-serving gossip as *bad*, and group-serving gossip as *good*.

Moreover, gossip, for social control not only works theoretically, but it can also be tested empirically. Wilson et al (2000) administered experiments about gossip that indicate people judge self-serving gossip as "highly inappropriate," but they consider gossiping to enforce social norms as appropriate -- when done responsibly. Additionally, failing to gossip may damage one's reputation (and therefore fitness within a group). Thus, gossip operates as a mechanism of social control, and when individuals misuse it, a group will punish them and decrease

their fitness, for example, by decreasing their status and access to resources. However, when used appropriately, gossip protects the group from cheaters, lowering the amount of self-serving behavior, and increasing the amount of group-serving behavior. In short, though the above scenarios are somewhat conceptual, and we need more tests to confirm Wilson et al's gossip experiment, all this still suggests that gossip serves altruism and GS by limiting self-serving behavior in groups.

Moreover, if gossip replaces grooming when group size increases (Dunbar, 1996), this implies that gossip can serve social control as group size increases. Small talk and gossip may oil the wheels of human relations by positively building camaraderie in the group, but social control gossip also builds camaraderie in a negative way, by serving as a protective mechanism that punishes self-serving behavior that damages camaraderie and group fitness. Since gossip in a group can simultaneously help many people hear about and also mark an individual's self-serving behavior as "bad," this acts as an efficient way of imposing social control -- where language usage through gossip asserts the group's norms and morals. In short, gossip increases the efficiency of social control as group size increases.

3.7 Language and the Transmission of Moral Codes

Besides the aforementioned functions of language that serve altruism and groups, language also enables groups to transmit a moral code from generation to generation.

A common language connects the members of a community into an information-sharing network with formidable collective powers. Anyone can benefit from the strokes of genius, lucky accidents, and trial and error wisdom accumulated by anyone else, present or past. (Pinker, 1994: 16)

This includes the accumulation of information that promotes altruism, and though Pinker discounts multi-level selection (Pinker and Bloom, 1990; Pinker, 2002), the passage above hints at how language might help facilitate group selection.

Language connects people in an information nexus that facilitates comprehensive cooperation, and though individuals can benefit from this accumulated information,

groups also benefit from it -- in competition with other groups. If so, then one group can out-compete another group based on their information resources, and this includes the kinds of altruistic (not necessarily ethical) resources they possess.

For example, regarding how language helps strengthen groups, the group with the best knowledge about how to get and produce food, wage war, and manage health and medicine would have an advantage to out-compete groups with poorer information. Moreover, groups with better information technology might have better access to the storage and retrieval of survival enhancing information. And groups might out-compete rivals if they possess high quality information and communicate it among themselves better than their rivals do. Hence, a group might decrease the fitness of a rival group by removing or weakening their ability to communicate with each other, by damaging or stealing their survival-enhancing information, or by taking away their linguistic identity -- forbidding the use of their language. In sum, language helps groups convey survival-enhancing information, including altruistic and moral information, and a group that communicates this information more effectively than its rivals gains a survival advantage.

The main questions here concern whether groups transmit *moral information* across generations and if this strengthens the collective power of a group in relation to rivals. For sure, groups do transmit moral information across generations. Jews, Muslims, Hindus, and Buddhists have verbal traditions that have transmitted moral and ethical teachings between generations. For example, in Hindu oral tradition...

The *Ramayana* encodes many of the cultural values of Hindu civilization: from a cosmology of cyclically-recurring eons, to a stratified social order and a patriarchal, extended-family structure based on arranged marriage, to the overarching theme of *Dharma* a central cultural concept suggested by terms like "morality," "duty," "cosmic order," or simply, "the Way." (Lutgendorf, 2004: ¶ 3)

With this in mind, it will be helpful to explore related issues concerning language, group identity, and how groups transmit moral information to regulate behavior.

3.8 Group Identity and Code Regulation of Group Behavior

The issue of language and group identity is large enough for a book length discussion (McKirnan and Hamayan, 1980). For our purposes, however, we need to reiterate that language can serve to mark group boundaries, as it helps insiders and outsiders identify each other (LePage and Tabouret-Keller, 1985). Moreover, as Crystal (1997) suggests, groups use language in numerous ways not just to provide information or content, but rather to assist in the expression of group identity. Chants at political rallies ("four more years!"), shouts of "amen" at church meetings, and phrases cheered at football matches (such as, "what an [expletive] home support!") all employ a language that "unites rather than informs" (Crystal, 1997: 13). Additionally, spoken or written language informs listeners or readers about the person who produces it, including their region, status, education, work, sex, and personality. Thus "a major function of language is the expression of personal identity -- the signaling of who we are and where we belong" (Crystal, 1997: 13).

Though this reiterates previous points, it also highlights an important nuance that language not only helps create groups, it also creates a group's abstract identity. At the same time, expressions of group identification also include implicit moral judgments about what speakers think is good and bad. Chanting "four more years!" at a political rally signals that one shares the values of a particular group and is willing to act on those values. Although we can deceptively identify ourselves with a group, other members will expect us to engage in certain actions: to back up our words with actions, to expend money to support the party, or even, depending on how outsiders perceive us, to risk social status. Additionally, in more volatile political contexts, publicly stating one's view can incur the ultimate sacrifice, so that group identification would not be worth cheating. Whatever sacrifices we incur by making these identification statements, benefits follow as the in-group accepts and protects members, for a group that protects its individuals will be stronger than a group that does not. In short, some of the important functions of language provide groups with a value-laden identity that helps distinguish them from other groups. Moreover, it is essential that language function like this for group-level

adaptations to occur that depend on higher-level processes and that facilitate altruism or morality.

In addition, language not only *identifies* groups and their members, it also helps to *regulate* the behavior of groups and their members. In fact, if the codes do not regulate group behavior, besides being meaningless, they could not influence group selection. Though previous sections have argued that language-based moral codes *can* regulate individual and group behavior, this section briefly shows that moral codes *do* regulate behavior. For example, Sober and Wilson randomly sampled twenty-five cultures from the Human Relation Area Files, and they conclude that...

The current evolutionary view of human behavior tends to portray individuals as free agents who can employ any strategy they want to maximize their inclusive fitness. This view does not deny the existence of social norms, but it does accord them a minor role in the *regulation* of behavior. In contrast, our survey suggests that human behavior is very tightly *regulated* by social norms in most cultures around the world. (Sober and Wilson, 1998: 165, italics mine)

Incidentally, because of this, Sober and Wilson claim that social norms serve as a mechanism "that substitutes for genealogical relatedness, making group selection a strong force in human evolution" (1998: 166).

To demonstrate their point, Sober and Wilson cite the example of the East African tribal struggle between the Nuer and Dinka. Nuer tribal norms apparently helped the smaller but well organized Nuer tribe to conquer the larger Dinka tribe. For example, the Nuer followed norms that guided how they divided the spoils of conquest and how they resolved internal conflicts, norms that the Dinka did not follow. In addition to Sober and Wilson, Brown (1991) also provides evidence that norms can regulate human behavior. He lists a number of cultural universals: inheritance rules, language employed to manipulate others, laws of rights and obligations, sanctions for crimes against the group, and basic moral sentiments. We know that groups put much time and effort into following these moral cultural universals, e.g., through religious rituals and acts of conflict resolution, showing the universal importance of norms regulating the life of human groups. Wilson (2002) also gives numerous in-depth examples (from Judaism, Calvinism, and

Tibetan Buddhism) showing how religious norms clearly regulate behavior and facilitate group selection.

Furthermore, the cultural universal of proverbs shows how humans use language to regulate behavior. Strungaru and Schiefenhövel (2002) discuss how cultures use proverbs couched in language to teach wisdom, altruism, and morality. Empirically measuring the practical effect of proverbial wisdom is difficult, but Brown and Paul (2002) define it as problem solving ability in complex, subtle, novel, or uncertain social situations. Thus, wisdom helps one apply social-moral information and expertise in daily life, and Brown and Paul argue that, in order to benefit from the wisdom communicated in proverbs, people need to understand and apply them to specific contexts.

Their research shows that diminished wisdom frequently occurs in subjects with agenesis of the corpus callosum, a brain disorder that inhibits the processing of information across both sides of the brain. This diminished wisdom refers to a decrease in the social problem solving ability just mentioned. The dysfunction in these patients, which comes from an *absence* of communication between brain hemispheres, suggests that the *presence* of such communication is required to process and apply proverbial wisdom. In a word, empirical studies show that normally developed humans can neurologically process moral information in proverbial language, information that regulates behavior. Nonetheless, despite all these examples of moral codes regulating behavior, we still need to demonstrate more fully how groups employ the *functions* of language to regulate behavior that facilitates *structural*-level (or group-level) adaptations. And before we do that, we need to consider the dark side of group selection as well.

3.9 The Dark Side of Group Selection

We know, from religious or military traditions for example, that groups teach groupishness. Paradoxically, the Parable of the Good Samaritan, which explicitly teaches the Golden Rule, also clearly criticizes those who only behave altruistically toward the in-group and not toward outsiders. That is, the story teaches disapproval

about moral behavior that *only looks out for its own*, (and it also promotes altruistic behavior towards the out-group, a topic to be discussed in the next chapter).

However, in-group-only altruism raises a serious problem, which we can call the dark side of group selection. Hartung (1995) discusses the topic of in-group versus out-group morality. His data on the subject comes from the ethical and historical writings of the Judeo-Christian traditions. Hartung argues, contrary to popular sentiment and the aforementioned story of the Good Samaritan, that Judeo-Christian morality provides no basis for altruistic behavior directed toward out-group members. On the contrary, Hartung claims that Judeo-Christian norms explicitly teach a full-blooded in-group morality. Hartung's argument is important because below I will argue that linguistically mediated moral codes can replace genetic relatedness creating a virtual relatedness that fosters altruism and group selection. Therefore, at this point we need to reveal the serious flaws in the common thinking that is represented in Hartung's empirical and theoretical argument against GS.

First, it is crucial to note how Hartung sets up the topic of group selection. Although he suggests that the moral principles "love your neighbor as yourself" and "do as you would be done to" were explicitly to be applied *to in-group members only*, he also unequivocally rejects group selection -- despite the fact that he interprets these codes as severely groupish. Hartung's claim is problematic not only because standard accounts of Judeo-Christian morality maintain that these moral norms were initially meant to have universal application, but also because there are technical reasons in evolutionary theory to conclude that an in-group morality should tend to promote group selection. More specifically, Hartung asserts that group selection cannot create a natural harmony between in-group members, that is, a *biologically based* altruistic orientation.

However, note that group selection, as a pluralistic, multi-level, structural-hierarchical theory of evolution, does not need to start at the biological level in evolving organisms. *Initially GS may not evolve organisms but groups, and it may select for socially controlled, rather than biologically controlled, behaviors.* Moreover, GS does not need to be responsible for a natural harmony between

members; other mechanisms may impose, facilitate, or create the degree of harmony that is necessary for groups to function, for example through linguistically based moral norms and enforcements. That is, GS does not need to facilitate morality, but morality can promote GS.

Consider the phenomenon of fierce punishments that enforce moral norms. We now see that even severe norms do not need to emerge from some puzzling natural kindness that was first produced by GS. On the contrary, strict norms might enable a group to coalesce and *after that* to become subject to GS pressures. GS does not have to evolve natural or biological niceness; other mechanisms -- some severe, some not so severe -- can evolve altruism. Perhaps, after many generations, the resulting altruism may be "biologized," that is, genetically coded and transmitted. The present argument neither requires nor denies that possibility. Instead, it emphasizes the role that *a linguistically based moral culture* might play in facilitating altruism, rather than how our biological make up might do so. Hence, I have focused on how groups structurally coalesce and evolve based on the functioning of linguistic moral mechanisms, subsequently becoming subject to GS pressures, rather than on how GS might evolve biologically-based altruism in individuals.

3.10 Introducing Virtual Relatedness

Hartung also discusses Hamilton's (1964) idea of relatedness, connecting it with altruism and the Golden Rule: *love your neighbor as yourself*. Hartung translates this rule into biological terms as *love your conspecific as if $r = 1$* . If the degree of relatedness (r) somehow became 100%, or $r = 1$, then there would be no genetic conflict of interest among group members, and they would therefore theoretically sacrifice themselves for others because they would share identical genetic interests -- like twins. Hartung concludes, however, that because self-sacrificial altruism opposes individual reproductive fitness, natural selection would eliminate it. That is, selection terminates self-sacrifice.

I argue, by contrast, that such altruism could exist *if there were mechanisms to maintain it and to protect the individuals who practice it*. As long as there are mechanisms that support altruism, that police against cheating and intruders -- and as long as one insists, with the group selectionists, that we compare between altruistic and selfish groups rather than merely between individuals -- the theoretical and empirical possibility of group selection for altruism clearly emerges. Hamilton's own work does not explicitly extend relatedness in this way. His rule states that altruism pays when b , the fitness *benefit* to an organism that receives help, multiplied by r , the *relatedness* of the receiver to the helper, are greater than c , the *cost* of the behavior to the helper or, in short, as long as $br > c$ (Hamilton, 1971). This kind of inclusive fitness does not necessarily entail a group-level structural adaptation. Instead, it simply predicts that genetically related individuals will behave altruistically toward each other. The less-related or non-related members of the group will not necessarily benefit from these altruistic acts, since kinship relations drive them.

In contrast to Hamilton's Rule, I would like to introduce the concept of *virtual relatedness*, which builds on Sober and Wilson's idea (1998: 166) that social norms, as a mechanism, can substitute for genealogical relatedness. Virtual relatedness, or vr , expresses how an individual treats another regardless of biological relatedness; that is, in Darwinian terms, vr reflects how *familiarly* one individual treats another individual regardless of actual kinship ties. Thus, from a biological perspective, vr quantifies *as-if-behavior*. Behave toward your unrelated neighbor *as if* you were related; treat her *as if* she were your kin, or even *as if* she were you. Take the case of a boating accident in which a teacher sacrifices her life to save her student; the teacher's genes are lost, and the unrelated student's genes live on. In this case the teacher gave 100% of her fitness to an unrelated individual; she acted *fully as if* the unrelated student were herself, without any interference from genetic conflict of interest. In this particular case, because there was a complete sacrifice of fitness, vr would be expressed as 1.

To elucidate vr , imagine a selection space with both related and unrelated individuals (see Figure 3.1 below). In this space, assume that all S individuals

(shown as diamonds) share a genetic relatedness of $r = 0.5$ and all A individuals (shown as circles) also share a genetic relatedness of $r = 0.5$, but that S and A individuals are completely genetically unrelated. Hence for S and A individuals, $r = 0$. This degree of genetic relatedness will predict altruism among related individuals; moreover, there also exists the possibility of reciprocal altruism between S and A individuals.

The concept of virtual relatedness suggests that behaviors completely separate from inclusive fitness and reciprocal altruism might occur in this world. Biologically, a S-A relationship would be expressed as $r = 0$ because they are genetically unrelated. If however an A individual were to behave altruistically and sacrificially toward an S individual, a form of virtual relatedness would emerge where, for example, vr could be 0.5. Hence, vr is a measure of behavioral relatedness, not biological relatedness. Moreover, it is not a function of a relationship between replicators; it is an indicator of altruistic behavior that ignores kinship.

Figure 3.1: Space of Biologically Related and Unrelated individuals

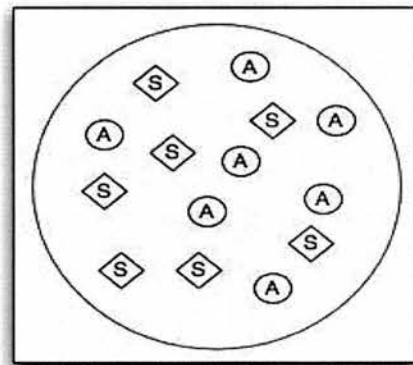


Figure 3.1. In this space, there are two groups of kin, the A group and the S group. S brothers biologically share a relationship of $r = 0.5$ and A brothers share a biological relationship of $r = 0.5$, but the biological relation between any S and A non-kin-pair is $r = 0$. A behavioral relationship between an S and A non-kin-pair could be virtually similar to a kin pair, however. This would be expressed as a measure of virtual relatedness, e.g. as $vr = 0.5$.

Imagine that an A individual has an adopted S individual as a brother and that he behaves towards him as if he were a biologically related brother. In this case, the

biological relatedness would be $r = 0$ but the virtual relatedness would be $vr = 0.5$. In addition, because vr concerns behavior, an interesting alternative scenario could play itself out. Imagine two A-group individuals who are biologically related brothers, where their biological relatedness is $r = 0.5$, but who have come to hate each other. In actual behavioral terms, their virtual relatedness could be $vr = 0$. In this case they are closely related biologically but behave *as if* they are not, since they show no altruism but rather spite toward each other.

The urgent question with regard to virtual relatedness is how it emerges. This is where the moral code comes into play. For the purposes of the discussion, let's imagine that the moral code is some version of the Golden Rule. We express this rule as g , where g represents any specific moral code, idea or belief couched in language that facilitates altruistic behavior. Now consider the case of the altruistic teacher, which we contemplated above. Presumably the teacher has internalized g through a very complex process, which was partly a linguistic one. We must note that this fact makes g linguistically detachable from the teacher; hence g is not identical to vr , which expresses the degree of as-if relatedness shown in her actual behavior. Yet one's linguistically describable moral code, g , clearly *influences* the degree of vr in those conspecifics who have learned and internalized it.

An interesting example of this phenomenon comes to light if we vary the adopted brother scenario above. The adopted brothers' level of biological relatedness is $r = 0$, but their behavioral relatedness could be $vr = 0.5$. Imagine that their behaviors reflect a vr of 0.5 because the two brothers falsely believe that they are genuinely biological brothers. Though brothers by adoption alone, they treat each other altruistically *as if* they were brothers (vr represents this altruistic behavior); and they behave altruistically to each other because they believe that they are brothers (g represents this belief). Their false belief that they are brothers serves as the mechanism that causes them to behave as if they were related.

The belief, or moral code (g) does not have to be false in order to increase vr . We could have a different belief, such as "treat your adopted brother as if he were your real brother." Again, note that vr represents the actual level of altruistic behavior, while g represents whatever linguistically mediated moral codes, ideas or beliefs

facilitate the altruistic behavior. *In short, the altruistic code, g , facilitates vr , the altruistic behavior.* In the case of the teacher, vr represents her sacrificial behavior on behalf of an unrelated student, and g expresses the moral code that she acquired through a complex process of acculturation in which language serves as an essential ingredient. For example, perhaps when she was still a child, the teacher's parents taught her in word and deed that *doing-as-you-would-be-done-to* is a good thing.

In short, basic empathy (Preston et al., 2002), linguistic argument structure, i.e., who does what to whom (Calvin and Bickerton, 2000), and the linguistically enabled capacity to categorize and recategorize, that is, to consider X as if it were Y (Chomsky, 1965; Monroe, 2002) all facilitate the creation, comprehension and application of altruistic principles (g) through language. Such principles (like the Golden Rule) can be both internalized codes and external codes that are enforced through social controls. Altruistic-moral codes of this type serve as a sort of social calculus, which serves to promote altruistic behavior (vr) towards unrelated individuals despite potential genetic conflicts of interest.

Seen from the perspective of group-level adaptation, g influences group members toward treating other members of their group as virtual kin. Thus, g acts as a lever to ratchet up levels of altruistic behavior. Through vr ratcheted and maintained by g , individuals engage in fitness-sacrificing behaviors: caring for the sick and less fit; giving food to the needy and weak; providing shelter for the poor; providing medicine, resources and hygiene for the disabled; and performing a variety of costly behaviors that are necessary to maintain and police a moral order. Although I have used the Golden Rule as a shorthand for such altruistic moral codes, much further work would need to be done to specify the specific behaviors that it entails. The central goal here has been merely to show how linguistic moral mechanisms (which represent a functional hierarchy) might influence group-level adaptations (which are expressed in structural hierarchies). For this purpose it was sufficient to work with the variable g and to use the Golden Rule as a short summary of what is potentially a much larger moral code.

Consider the results we have achieved in terms of two groups of equally distributed biological membership, Group A and Group S, in which altruistic behaviors and selfish behaviors (respectively) dominate. Group A, which is able to increase vr by g , would be more reproductively fit at a group level than Group S, which is not as successful at increasing vr by g . That is, the moral code of Group A enables its members to behaviorally treat non-kin *as if* they were biologically related. Moreover, this selective advantage will remain though the average *individual* fitness of A-members becomes lower than that of S-members because of their sacrificing individual fitness for the sake of the group. This simple scenario avoids the averaging fallacy -- that is, comparing fitness by means of individual fitness across groups instead of comparing fitness between groups (Sober and Wilson, 1998) -- because it compares groups and not individuals. The key difference between the two groups is the manner in which Group A (the more altruistic group) increases vr , the expression of virtual relatedness. Since both groups will have some kind of linguistically based moral code, the fitter of the two groups is the one that is able to *ratchet and maintain* virtual relatedness more efficiently than the less fit group.

3.11 The Reciprocity Principle

It is significant that g *plus* vr , which in theory can help facilitate individual altruism and group fitness, finds its linguistic formulation in the reciprocity principle, "Do as you would be done to" -- or, as George W. Bush puts it, "We must all hear the universal call to like your neighbor just like you like to be liked yourself" (Weisberg, 2001: 28). Taking language usage as a guide or an empirical pointer, we can track the rule of vr as it is articulated in a variety of cultures. Such cross-cultural studies are interesting for at least two reasons. (1) Stating the rule requires the use of the argument structure *who does what to whom*, a foundational element of linguistic communication and universal grammar, so that the Golden Rule requires complex human language for its expression. (2) That the Golden Rule finds linguistic expression in cultures and languages all around the world strongly supports the idea that g exists as a cultural universal, a linguistic artifact which

groups use to ratchet and maintain *vr*. Consider some of the examples collected by (Wikipedia, 2003):

900 to 500 BCE—Judaism

"You shall not take vengeance, nor bear any grudge against the sons of your people, but you shall love your neighbor as yourself."

700 BCE—Zoroastrianism

"That nature only is good when it shall not do unto another whatever is not good for its own self."

500 BCE—Buddhism

"Hurt not others in ways that you yourself would find hurtful."

500 BCE—Confucianism

"Do not unto others what you would not have them do unto you."

400 BCE— Socrates

"Do not do to others what would anger you if done to you by others."

150s BCE -- Brahmanism and Hinduism

"This is the sum of duty: Do naught unto others which would cause you pain if done to you."

58 CE—Christianity

"And as ye would that men should do to you, do ye also to them likewise."

"Love your neighbor as yourself."

90 CE—Epictetus

"What you would avoid suffering yourself, seek not to impose on others."

800 CE—Islam

"No one of you is a believer until he desires for his brother that which he desires for himself."

The rule of *vr*, as linguistically articulated in the Golden Rule, is an expression of the effort of human groups to get members to treat non-kin *virtually as* kin, or even as more than kin; it involves the attempt to create a behavioral relatedness where biological relatedness does not exist. Incidentally, though we call this rule the ethic of reciprocity, it is different from reciprocal altruism in one significant sense. It is not simply saying to return to others the good (or bad) they have given you. Rather it says: behave towards others as you would have them behave towards you, or regard others *as if they were you* -- or, in Darwinian terms, treat others as if they shared

100% of your genes. Treat your neighbor as your twin. Care for conspecifics as clones. Hence, it attempts to push individuals not only to reciprocation, but also to a virtual form of the behavior that facilitates inclusive fitness and kin selection, except that the recipients of the altruistic benefits may be biologically related or unrelated. That is, it calls people to the same kinds of behaviors that we associate with kin selection, behaviors that are not conditional on reciprocation. Again: behave as if you were genetically related to this person, even though you are not.

3.12 Group Selection, Morality and Religion

Our focus here has been not so much on how this structural-functional hierarchy would create moral individuals through processes of biological selection, but much more on how a linguistic moral mechanism could at the functional or behavioral level increase the virtual relatedness of individuals. These individuals would in turn coalesce structurally into a group of altruists, especially when compared with another group, which is less altruistic. The present thesis thus involves a significantly different claim than Sober and Wilson (1998) whose view of group selection holds that genes underlie altruism. By contrast, I have argued that, if a *g*-code facilitates *vr* behavior in a group that coalesces and remains more altruistic than another group, the altruistic group will outcompete the selfish group. Thus, group selection will occur because the altruists out-survive the egoists, regardless of what gets biologized. An altruistic cultural influence in *g*, producing a behavioral element in *vr*, will help the group attain a higher level of populational or biological fitness than a group with a lower level of *vr*.

In this scenario, it is important to recognize that the processes, mechanism, and behaviors affecting altruism can be pluralistic. Some of the altruistic behavior can be produced by inclusive fitness and by reciprocity; not all altruism will be the product of linguistic moral mechanisms or the moral code. In the end, from the structural or group-selection perspective, a crucial factor is what happens to the group in comparison with other groups.

Moreover, group selection may not seem very moral. Thus, for example, Hartung (1995) and Pinker (2002) decry group selection as involving the morally repugnant

kind of evolution co-opted by the Nazis, which is unlikely to produce *good-natured* behavior. Although this is an important concern, the primary focus of my argument is *descriptive*: to discover whether group selection actually occurred during human moral evolution. It may well be that we will judge much of the inter-group conflict that accompanies group selection as morally wrong. But whether or not GS has a dark side is beside the point, at least at first; we simply want to know *whether* it happened. Moreover, there are ways around the dark side of group selection. Moral judgments concerning group selection depend on how the moralizers define the group, and since groups manage their moral mechanisms, it should be possible to expand the circle of the group (Singer, 1981; 1997). I return to this issue below and in greater detail in the following chapter.

This brings us to a remarkable feature of Hartung's discourse about inter-group conflict: his dismissal of group selection may actually exemplify a GS process. Hartung contends that Judeo-Christian teaching has explicitly promoted an in-group morality at the expense of out-groups. At one level, this seems to be true: according to the historical data of the Jews, the ancient Israelites in their conquest of Canaan appear to have dispossessed at least seven nations. Moreover, the historical data describes the manner in which, in the process of their violent conquest of Canaan, they destroyed out-group members. Moreover, the records provide other indicators that group-level selection was involved in this process.

Our particular interest lies in the role of the Judaic moral code in relation to the groups involved, and particularly how it defines in- and out-groups. Jews verbally prescribed circumcision as a group marker, and they developed numerous strict laws concerning religious ceremonial observance within the group. At the same time, they had a moral law that we can summarize as $r = 1$ or "love your neighbor as your self." That is, Commandments 5-10 generally express the logic of the Golden Rule in human relations. In simple terms, people do not like being stolen from, so they are told not steal from others; people do not like being killed, so they are told not to kill, and people do not like having adultery committed against them, so they are told not to commit adultery.

Additionally, many of the punishments for failure to observe the law were extremely costly to the unlawful, including in many cases the death penalty. It is important to note, however, what emerges when one evaluates the way the Hebrews treated unlawful individuals in their own in-group. *They explicitly treated them as outsiders.* That is, those who defied or broke the code -- offspring, parents, siblings, identical twins, cousins, reciprocal neighbors, and any kin or non-kin -- were to be treated as outsiders. This is especially true depending on the seriousness of their offences against the moral mechanism, and this is significant because it shows the way the moral code defined the group. Egoistic behavior that decreased virtual relatedness was not tolerated, and to a significant degree it was *equally* not tolerated in both the in-group and the out-group. Failure to adhere to the moral code, at least in principle, made anyone -- regardless of relatedness -- an outsider.

If we assume that this severe rule of the moral code actually operated in Hebrew society as depicted in the historical data, then these punishments may represent examples of what some call the violent side of group selection. But even if the outsider principle was not actually applied to in-group members in so severe a form, the process described in the texts may still provide a model of the structural-functional process of virtual relatedness. This is because *in either case* the moral rules served to define the group: functioning in this way, they increased the group's fitness by ratcheting the virtual relatedness of individuals. One may wish to raise moral questions about the fact that this group, with its severe moral code, dispossessed numerous groups with apparently less strict in-group altruistic codes. But the moral judgments remain secondary to the question of whether linguistically mediated *vr* actually plays that role that it appears to play in supporting altruism.

Wilson (2002) strongly argues for the group selective nature of Judaism (and other religious traditions), which we cannot review here in detail. He claims the Ten Commandments and the Golden Rule facilitate group-level adaptations; moreover, the Torah commands, "be fruitful and multiply." That is, this moral code teaches its group to be biologically successful. Wilson also claims that Judaism instructs its members "to behave very differently among themselves than toward members of other groups" (2002: 135). However, this is only partly true, and if it were

completely accurate, then it would contradict my assertion above that the code defined the group regardless of relatedness. Wilson argues that multi-level selection theory predicts a hypocritical moral code that simultaneously preaches to love one's neighbor but to destroy threatening genetic outsiders. However, my claim emphasizes that the code also idealistically defines outsiders by their behavior. For example, the Torah commands the death penalty *for insider Jews* for murder, adultery, incest, homosexuality, bestiality, cursing one's parents, human trafficking, Sabbath breaking, idolatry, and blasphemy. Jews may have behaved differently to genetic outsiders, but their moral code also caused them to treat genetic insiders who broke the code *as if they were outsiders*, thus making the moral code a defining factor of the group.

In short, GS must cause a group to coalesce to the point of becoming a virtual organism, and organisms build up immunities to viral invaders, which block their selfish influence on the unity of the body. Hence, if the moral code defines the group and destroys influences that would decrease the *vr* of members, this is the kind of code that would facilitate group selection -- no matter how repugnant we may find it. In summary, it appears that there is strong evidence that linguistically based moral mechanisms are able to advance virtual relatedness, to punish offenders (by decreasing their fitness) or to eliminate them (by selecting them out of the environment), regardless of their actual biological relationship or status as insiders or outsiders. The code not only characterizes the group; *the code literally identifies, perpetuates, and regulates the group.*

3.13 Conclusions

This chapter demonstrated how language enables individuals to coalesce into groups, how language can help facilitate altruism, especially between non-kin within a group, and finally how a group unified and made altruistic with the help of language might become subject to group selection pressures. The increasing acceptance of GS supports this thesis, with strong theoretical backing from Sober and Wilson (1998) and diverse empirical examples (Dicks, 2000; Bekoff, 2002; Wilson, 2002). Nevertheless, we still need more empirical support.

We still need somehow to figure out how to test rigorously extant ideas about levels of selection -- group selection 'versus' individual selection -- and studies of the evolution of social morality are good places to focus for expanding our view. (Bekoff, 2002: 86)

Fortunately, regarding human sociality and its relation to morality and language, we are looking in a good place; therefore, further research seems promising. Nevertheless, in spite of the potential of studies on the relation between language, morality, and multi-level selection, there is still something unsettling about group selection. That is, group selection has racist and other unethical implications. To quote Pinker on the subject:

Group selection, in any case does not deserve its feel-good reputation. Whether or not it endowed us with generosity toward the members of our group, it would certainly have endowed us with a hatred of the members of *other* groups, because it favors whatever traits lead one group to prevail over its rivals. (Recall that group selection was the version of Darwinism that got twisted into Nazism.) This does not mean that group selection is incorrect, only that subscribing to a scientific theory for its apparent political palatability can backfire. (Pinker, 2002: 259)

A group may be altruistic toward insiders but extremely selfish to outsiders. Moreover, if a group thinks they are genetically, socially, or morally better than another group, and if they think outsiders might corrupt their gene pool, then they might use GS to justify racism. Unfortunately, if group selection has taken place in this way, it is simply an ugly and evil fact; however, the discussion on language, evolution, and morality need not stop here. As previously mentioned, a golden-rule type code (*g*) based in language enables us to re-categorize within group non-kin as if they were kin and therefore behave towards them as if they were kin -- through *vr*. If this is true, then *g* not only can facilitate *vr* towards within-group non-kin, but it could also do so for out-group non-kin, and this takes us beyond the ethically negative implications of GS and into the focus of the next chapter.

CHAPTER 4

LANGUAGE AND MORALITY BEYOND GROUP SELECTION

4.1 Real Altruism Directed to Out-group Members

The previous chapter discussed the role language plays in facilitating altruism, groupishness, and multi-level selection. In this process a moral code g , encourages virtual relatedness vr , where individuals treat in-group non-kin as kin. This transcends reciprocal altruism because it encourages the same kind of altruistic behaviors associated with kin selection, behaviors that are not conditional upon reciprocation. Hence, vr is *not* expressed as "treat only your fellow reciprocators as if they were you," but rather, "treat all your neighbors as yourself." Of course we can debate the definition of neighbor, but for this discussion it refers to non-kin in one's group. Thus, if individuals express vr to non-kin within a group, where they sacrifice their fitness for the group and where groups selfishly compete against each other, questions naturally arise regarding vr directed toward out-group members. Hence, this chapter focuses on how virtual relatedness might explain how individuals and groups might direct and maintain altruism towards non-kin and out-groups.

Concerning extant altruistic behavior directed to out-group non-kin, it is interesting to consider how the selfish gene perspective interprets it. First of all, *any altruistic behavior* directed towards out-group members exceeds the explanatory confines of inclusive fitness because out-group members generally do not comprise kin. Of course, socially ostracized kin members would represent a special example of genetically related individuals who are also out-group members; however, in

almost all cases these outsiders are also non-kin. Therefore, in a word, inclusive fitness cannot explain altruism to non-kin out-group members.

Secondly, and in contrast to inclusive fitness, reciprocal altruism may produce behavior directed to non-kin out-group members that still meshes with selfish gene theory. However, though reciprocity exists between genetically unrelated individuals within groups as a cultural universal (Brown, 1991), the theory of reciprocal altruism (Trivers, 1971) does not explain *vr* to in-group or out-group non-kin where reciprocity is not a condition of the altruism. That is, reciprocal altruism does not explain *kin-like behavior directed at non-kin*. The following chart depicts the coverage of these theories of altruism, where KS is kin selection, RA is reciprocal altruism, GS is group selection, and VR is virtual relatedness.

Figure 4.1: Coverage of Theories of Altruism

Relation	KS	RA	GS	VR
Altruism to kin	Yes	Yes	NA	NA
Altruism to in-group non-kin	No	Yes	Yes	Yes
Altruism to out-group non-kin	No	Yes	No	Yes
Altruism to non-reciprocal out-group non-kin	No	No	No	Yes

Related to these theories of altruism, selfish gene theory (SGT) often *explains away* radical altruism, seeing it as *apparent altruism* that is actually selfish, or SGT may see selfish reciprocity in what only *appears to be* radically altruistic behavior. Most importantly, if the altruist gets no reciprocal benefit, SGT forces us to understand the altruist's behavior as maladaptive, and/or manipulated by another selfish conspecific, which inevitably results from a purely egoistic process. Nevertheless, even if we grant that some forms of radical altruism are maladaptive or manipulated, this does not mean that they all are. Hence, with these things in mind, we will set the stage to discuss morality and altruism that may transcend GS and SGT by summarizing some kinds of *apparent altruism* below. This will exemplify ways in which SGT construes -- and sometimes possibly misconstrues -- altruistic behavior as selfish, when the possibility remains the altruism may be real.

- **Genetic Lag:** The environment changes faster than genes, so behaviors lag and match former environments. Thus, behaviors that increased fitness in the environment of evolutionary adaptation (EEA) do not increase fitness in

current conditions. This is anachronistic fitness. Adoption forms a maladaptive expression of parental care in modern society where kinship ties are weak, but adoption increases genetic fitness in smaller (EEA type) communities where kinship ties are stronger.

- **Phenotypic Plasticity:** Phenotypic behavior changes in response to the environment, but the response is maladaptive. This complements genetic lag but proposes *not* "that we have unchanged (vestigial) behaviors that fail to keep up with environmental change but that environmental change induces change in behavioral phenotypes (Post et al., 2002: 221). Parents may invest in an adopted and genetically unrelated child (behavior not predicted by SGT) because the cultural environment shows negative attitudes about distressful childbirth.
- **Pleiotropy:** A gene determines two or more apparently unrelated characteristics in an organism. Thus, radical altruism maladaptively emanates from the adaptive traits of a related genotype or phenotype. Altruistic adoption, though seen as maladaptive in SGT, may connect to genes for bonding with one's progeny.
- **Manipulation:** The egoist manipulates behavior from the altruist. For example, in nest parasitism, cowbirds manipulate the parental instincts of warblers and get them to sacrifice their reproductive success and invest it in cowbird offspring.

4.2 Group Selection is Selfish Selection Between Groups

Besides the fact that SGT still sees radically altruistic behaviors as selfish, we must also note that GS simply translates this selfishness up to a higher level. A group may be a collection of related and unrelated individuals, who by virtue of shared language, culture, values, behaviors, and history coalesce into a unit that competes against similar groups. However, though groups limit the selfishness of individuals and increase their altruism (through linguistically based moral codes and social controls), they still form a selfish group in relation to other groups. Even regarding reciprocal altruism, as it meshes with SGT, non-group and non-kin members exist at varying degrees of genetic difference and distance from one's in-group. Thus, for example, my in-group will experience riskier forms of reciprocity towards out-groups, depending on their genetic and cultural distance from my in-group. Therefore, it is likely that greater genetic, cultural, and linguistic distance between groups will decrease the reciprocal altruism between them. In other words,

reciprocal altruism may decrease as degrees of genetic, cultural, and linguistic difference increase.

Thus, it seems likely that individuals would tend to reciprocate with non-kin members of their own ethnic group for various reasons. These individuals would find non-kin in-group members more trustworthy for having the same language, values and norms, and because they would tend to share a closer genetic make-up than members of out-groups who speak different languages and practice different moral customs. More importantly, though SGT may interpret continued and non-reciprocated altruism to any out-group non-kin as maladaptive or manipulated, GS will also interpret such behavior as anomalous as well. That is, if we keep redefining the boundaries of the group to include previously unaffiliated non-group members in order to reciprocate with them or increase *vr* with them, our group would change or dissolve and cease to avail itself to GS pressures. In any case, GS does not predict and cannot produce universal brotherhood (Wilson, 2002).

Moreover, if we help members of a group with whom we compete, then at worst we might contribute to our own group's de-selection, or at best to the selection of two groups, if that is possible. Either one of these results appears to raise problems for group selection unless we label such altruistic behavior as maladaptive or manipulated. However, at least in many modern societies, we consider it virtuous and morally imperative to help outsiders who cannot reciprocate. Therefore, regarding these problems, two issues stand out. First, as previously mentioned, SGT and GS are not mutually exclusive, so we are not bound to an "either/or" perspective when looking at these problems. Hence, we may pluralistically employ both theories to help us resolve some quandaries. Second, though SGT and GS may interpret non-reciprocated altruism to non-kin out-group members as maladaptive, some forms of radical altruism may still transcend the limits of the explanatory power of selfish gene theory and group selection.

Keeping this background in mind, the second of these two factors stands as the theoretical foundation of this chapter, which is summarized below.

- Altruism directed to non-kin, non-group individuals or groups might not necessarily prove maladaptive in all cases.
- If so, these behaviors would exceed the explanatory limits of selfish gene and group selection theory.
- The theory of virtual relatedness helps explain how altruism can be directed and maintained to non-kin individuals and out-groups.

4.3 Atypical Altruism Transcending GS and SGT

Schloss (1998; 2002) has delineated numerous behaviors that may contradict selfish gene theory -- unless we label them as maladaptive. Here is a partial list that relates to the focus of this section:

- Heroic and one-time efforts to save others -- across kindred and group lines.
- Sacrificial philanthropy -- across kindred and group lines.
- Blood and organ donation across kindred and group lines.

Though these behaviors are rare, they are not thoroughly uncommon. Thus, a main question pertains to whether such actions simply misapply a core adaptive foundation or whether they represent higher-level adaptations, perhaps resulting from cultural evolutionary processes over and above biological adaptations.

Boyd and Richerson (1997) explain how cultural evolutionary processes enable larger groups to transcend selfish genes that keep behaviors within the bounds of kin altruism and reciprocal altruism that exist in smaller groups. These larger groups would experience cultural group selection, employing numerous mechanisms based in symbolic group demarcation that would increase altruism toward non-kin or to those who for various reasons cannot reciprocate. This *inability* to reciprocate could result from group size where the need to cooperate would cause members to behave altruistically towards other members that they may only meet once. Otherwise, non-reciprocity would result from a lack of altruistic ability in the beneficiary, for example, stemming from a physical disability.

Whatever the case, as a cultural evolutionary process including virtual relatedness enables individual group members to behave in ways that transcend biological evolutionary imperatives, both in action and in pace, the possibility clearly arises

for behaviors to emerge that surpass SGT and GS constraints. Moreover, even if GS and SGT perspectives label this radical altruism as apparent or maladaptive, this does not necessarily explain what kind of process or mechanism causes this kind of atypical behavior to emerge and function in the first place. Therefore, in this chapter, I will not only focus on mechanisms and processes that may produce these behaviors, but I will also claim that such behaviors do not always merit a maladaptive interpretation.

First, looking at the maladaptivity question, we may benefit from evaluating the particular kind of behaviors in question. How might any form of sacrificial altruism across kindred, reciprocal, and group lines function in an adaptive sense? If a Mother Teresa altruist (Ruse, 1991) sacrifices her individual fitness for out-group untouchables who cannot reciprocate, and if she teaches her in-group members to do the same, what adaptive benefit could this behavior possibly have? Biologically speaking, such an altruist would experience no fitness benefit from such behavior because self-sacrificial charity contradicts biological natural selection: her altruistic behavior ends her genetic future. Moreover, such "Mother Teresa" altruists tend to practice celibacy. Thus, absolutely no reproductive fitness gain is attained, only cost.

However, since the mechanisms that produce this kind of behavior are primarily cultural, and since such behavior transcends biological imperatives, is it possible that any higher-level cultural adaptation comes from this kind of behavior? Taking a group perspective, the recipients of the altruism would clearly receive fitness benefits; hence, though the sacrificial behavior could be considered maladaptive for the altruists, it is clearly adaptive for the beneficiaries. For example, a "Mother Teresa" altruist could sacrifice her fitness for 100 people, and because of her efforts it is possible that each one of these individuals could increase their reproductive fitness. In this radical case, a Mother Teresa altruist serves as a kind of fitness agent for others, and although she is reproductively unfit in a biological sense, she may pass on her cultural fitness to others. Moreover, this may actually be the case since there are approximately 70 societies of charity founded on the "cultural" system of Mother Teresa. Upon closer scrutiny, these groups may turn

out to be egotistical; however, for the purposes of this discussion, we will take seriously the hypothesis that they are genuinely altruistic groups.

Before moving on, we need to define cultural fitness. First, cultural fitness refers to the reproductive success of beliefs, norms, or morals that create, define, and propagate a group. Group membership might experience turnover, but if these beliefs, norms, and morals possess traits that cause them to reproduce in individuals that form a group, then these *ideas* will express a level of cultural fitness. (The last chapter of this thesis discusses a similar idea of memetic fitness). Second, this *group* will express a level of cultural fitness as well because the group becomes the vehicle in which a set of ideas gets reproduced and because these ideas might positively influence the group. Thus, cultural fitness has two components: (1) the cultural fitness of the ideas themselves; the ideas get reproduced in a group; (2) the cultural fitness of the group; the group benefits from its beliefs and norms. For example, a group may have cultural beliefs about honesty, hard work, and education that help it prosper economically, or it may have explicit rules about sanitation that keep members healthy. Incidentally, this two-fold expression of cultural fitness would not be possible without language, which enables a group to reproduce and transmit its ideas in ways that may positively influence the behavior of its members.

4.4 TFT Reciprocity Between Groups and Radical Altruism

Moreover, we must consider the multiple groups that interact in this process and the other adaptive benefits that might emerge. For example, the altruistic group does not receive fitness benefits from its beneficiaries; moreover, because of the nature of their activities, such altruistic groups require the support of other altruistic groups (or semi-altruistic groups) to exist. Hence, the altruistic group must receive aid from at least one support group, and though the altruists cannot reciprocate this assistance directly, it is still possible for the support group to receive some kind of cultural fitness benefit from supporting the altruists. For example, if it becomes known that the support group is assisting the sacrificial

altruists, the support group may win status, good reputation, or trust that helps it reciprocate between other groups.

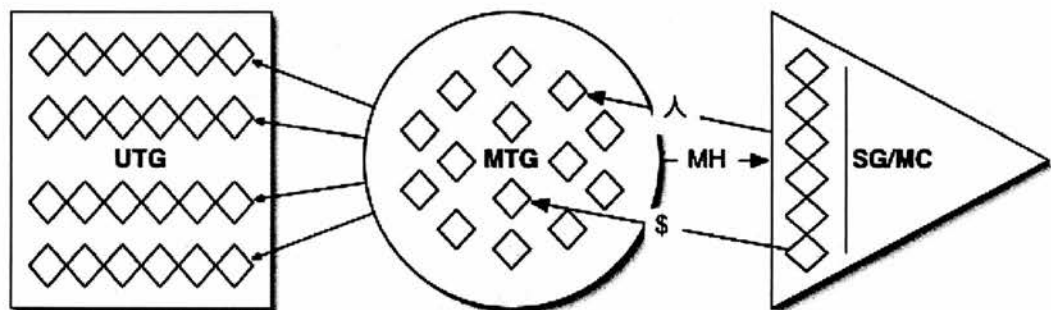
Hence, in this process, the existence of the sacrificial group requires and perhaps causes a reciprocal matrix to emerge between various groups. In this way, an altruistic Group A can receive *cultural fitness benefits* from a support group that help sustain Group A and its radical altruism. The support group provides for Group A so that it can aid the beneficiary group. Moreover, the beneficiary group can receive biological fitness benefits, through increased health, for example, and theoretically individuals in this group can receive a cultural fitness "lift" through education or other resources received from the actions of the altruists. This scenario at least theoretically demonstrates the adaptive potential of radical altruism from a cultural evolutionary perspective when we consider the interdependencies and cultural fitness benefits conferred on the various groups involved in this process.

The following description and diagram (Figure 4.2) adds some detail to this picture. Imagine three groups: an untouchable group (UTG), a Mother Teresa Group (MTG), and a support group (SG). Let's say that in the UTG, each diamond represents 100 people, so the 24 diamonds represent 2400 people. These people could suffer from any number of physical or cultural diseases, such as leprosy, STD's, alcoholism, unemployment, or criminality. Among other things, these problems will cause a lack of biological and cultural fitness. In the MTG there are about 12-14 radically altruistic individuals working with this particular UTG at any one time. The individuals in the MTG may change over time, but the group generally maintains its size and cultural identity. This means that there are about 200 UTG members for every 1 radical altruist from the MTG. The problem for the altruists in this situation concerns how to maintain their altruism while keeping cultural fitness as a group.

The support group (SG) from the main culture (MC) solves this problem and provides 2 kinds of resources to the MTG: (1) physical resources like food and money, and (2) human resources to replace the turnover and lack of biological fitness of the altruists. Hence, for this MTG, the support group (SG) consists of 6

sub-groups that provide human and physical resources. The dollar sign (\$) represents physical resources, and the Chinese character for person represents (!) human resources. To continue getting resources, some members of the MTG group mobilize help (MH) from the SG; this is represented by the (MH) in the diagram. That is, some MTG altruists mobilize human and physical resources from the SG so that the MTG can maintain cultural fitness. These mobilizers will use various means to motivate people to join the SG, including guilt, empathy, or religious commandments. Specifically, say that each altruist requires \$1,000 for monthly living expenses, totaling \$14,000 per month for 14 altruists. Also the MTG has a \$6,000 monthly budget, most of which is spent directly on altruism: giving medicine, food, or education to the UTG. This makes this MTG's monthly budget \$14,000 per month for a staff of altruists and \$6,000 per month for altruistic resources, totaling \$20,000 per month and \$240,000 per year.

Figure 4.2: Diagram of Support for Maintaining Radical Altruism



This may be an unrealistic budget in how it speculatively forecasts what might work in a given situation. However, it is not unrealistic in form or in the amount of money that real organizations actually raise for charitable purposes. For example, one organization, World Vision International, raised 1.2 Billion dollars in 2003, and it follows a similar (but more complex) model to the one presented here. Out of World Vision's budget, 9.8% went to fundraising, mobilizing help (MH) in our model, and 8% went to administration; 1.4% went to community education and advocacy (MH), and 80% went to humanitarian programs -- that is, MTG activities. Approximately 820 million dollars of this budget came from cash

donations, and another 435 million dollars came from "gifts in kind" (Vision, 2003).

In sum, altruistic organizations can actualize this kind of model, and in fact this is generally how many of them are managed. Thus, this model works in the real world, not only supporting altruists, but also providing the means for their altruistic work. Lastly, this model cannot work without language. The altruists employ language to teach cultural fitness to members of the UTG, and altruists also verbally mobilize help from the SG. That is, language plays an essential role in helping raise the cultural and biological fitness of the UTG and in mobilizing help (MH) from support groups (SG) so that the MTG can function altruistically.

The previous scenario concerns Mother Teresa-like altruistic individuals who sacrifice their fitness by serving in a *radically altruistic group*. This group requires the material support of *reciprocally altruistic groups* (SG's) or individuals whose altruism may experience reciprocity for the status or other cultural benefits they receive for their generosity. Incidentally, this model does not deny the possibility that some of the supporters may be acting, at least sometimes, in a purely altruistic way, but that is not the focus of this discussion.

This complex altruistic interaction benefits a *non-reciprocal or an a-reciprocal* UTG that receives both biological and cultural fitness benefits, either for individuals or for the whole group. In short, this *groupish* scenario elucidates the possibility of how radical altruism can express an adaptive edge. As a caveat, this scenario describes what can happen when a surplus of resources exists, and it does not attempt to explain how such altruism might emerge when there is no surplus. Nevertheless, we have here a form of altruism that transcends selfish group selection because it benefits non-reciprocal out-groups. Moreover, this scenario may still only be possible in a groupish sense, and human groups may possess the only mechanisms -- such as language -- to help selfish individuals overcome their Darwinian biological imperatives to the degree that produces this kind of radical altruism.

Moreover, a number of interesting ramifications emerge from this scenario. First of all, as already mentioned, it is possible to see how reciprocal altruism between groups can mirror reciprocity between individuals. As individualistic TFT altruism brings biological fitness benefits to interacting individuals, groupish TFT altruism confers biological and cultural fitness benefits to interacting groups. For example, rival political parties can make alliances that bring mutual benefit, or countries that rival each other can make alliances that bring mutual benefit, thus showing TFT at a group level. Thus, the whole idea of reciprocal altruism between groups raises an interesting research agenda within group selection studies.

Secondly, it is interesting to see the role that radically altruistic groups play within groups and between groups from a cultural evolutionary perspective. For example, a radically altruistic sub-group could exist within a group. To use a biological example, the radically altruistic sub-group could mimic white blood cells in the human body, protecting the main group from infection, containing and healing sickness, and providing immunity from disease. This radical altruism within a group could help increase the biological and cultural fitness benefits of the group. The scenario above already outlined how radical altruism might benefit out-groups. Though radical altruists would risk fitness, their sacrifice would benefit the whole group by increasing the biological fitness of weak members, and by teaching cultural fitness to those who fail to contribute to the group's cultural fitness.

Hence, biological and cultural fitness benefits could result from the actions of a sub-group of radical altruists in a larger group. However, as just mentioned, such radically altruistic sub-groups not only serve their own in-group, but they also serve out-groups, and to make matters even more complex small groups of radical altruists can even consist of members of unrelated out-groups. For example, a group of radical altruists in India might come from India, Poland, the US, Scotland, Sri Lanka, and Singapore. Nevertheless, it is possible that these groups of radical altruists may help facilitate reciprocal altruism between groups -- showing that radical altruism may require cultural reciprocity between groups.

Though we need to explore these scenarios more fully, this discussion shows how MTG's actually operate in the world. Moreover, researchers need to seriously

consider such scenarios because they may deal in real altruism and therefore do not deserve the label as apparent, maladaptive, or manipulated altruism. In other words, this approach addresses the potential failure of the SGT project to account for the complexity and potential reality of radical altruism in the real world. More importantly, as I will elaborate below, if these scenarios actually exhibit real altruism, then we may have found a factual way to avoid the moral nonrealism of Ruse (2002) and other selfish selectionists.

In some cases, the selfish shoe may fit; however, it seems overly cynical to call such altruism selfish in all cases, especially because cultural groups highly esteem radical altruism from their members -- or even when they see it in out-groups. Moreover, the adaptive picture changes when the researchers consider higher levels of order, such as cultural fitness and TFT reciprocity between groups that may partly support radical altruists. In addition, it is an empirical question to evaluate how TFT-between-group-reciprocity might, not only benefit cultural fitness, but also individual biological fitness as well. The biological world expresses this kind of inter-dependency; it is called the ecosystem, a group of interdependent life forms that inhabits and depends on a given environment. Hence, we may also need to consider cultural ecosystems when trying to understand group selection, cultural fitness, and the role that radically altruistic groups play in such inter-dependencies.

4.5 Radical Altruism Beyond TFT Between Group Reciprocity

In addition to the cultural ecosystem that supports TFT altruism between groups, an individual radical altruism also appears to exist. This radical altruism serves out-group non-kin, and the individual altruist may not belong to a Mother Theresa Group that uses cultural pressure to foster radical altruism. This altruism, for example, might find expression in individual heroic or sacrificial philanthropy across group and kindred lines. In such cases, individualistic radical altruism does not produce biological or cultural fitness returns, giving no fitness benefit, direct or indirect to the altruist. Moreover, regarding this altruism, researchers cannot easily detect forms of reciprocity occurring at different levels of order; therefore, it appears that purely sacrificial and non-reciprocal altruism exists at the individual

level. Though selfish gene theory must label such behavior as maladaptive, this does not actually explain it. Additionally, another theoretical perspective may lift the behavior out of the maladaptive realm and actually explain it as well. Thus, the researchers' problem mainly concerns theory choice for how we will interpret this kind of behavior.

One important example of radical altruism comes from atypical individuals or families who risked all to save Jews from being taken to Nazi concentration camps. Some of these altruists may have acted in line with the influences of a radically altruistic group and its cultural pressures, and others may have acted more on individual impulses. In all cases, however, the risks outweighed all biological fitness benefits, such as reproductive fitness to offspring. Moreover, cultural fitness benefits, status, or access to greater resources would be denied and not gained from such behavior, for the objects of altruism were generally non-kin and cultural outcasts. Stigma and not status were associated with helping such outsiders. Therefore, it seems that human behavior exists -- though uncommon -- that conforms to no clear evolutionary imperative. Namely, selection at the individual and group level cannot produce this kind of altruism because it confers no fitness returns to individuals or groups who do it. This behavior benefits non-reciprocal, non-kin, outsiders. Thus, this extreme altruism could only at best come as a by-product of features developed by natural selection. That is, a selectionist scenario cannot account for extreme forms of altruism unless we explain them as exaptations from pre-existing selected features.

With these issues in mind, we move the discussion to a level beyond how language simply interacts with altruism and morality to where it may affect radical non-reciprocal altruism to out-group non-kin. Thus, even if selectionists label this kind of altruism as maladaptive or manipulated, this still does not explain how such behavior emerges. That is, if we claim real altruism exists as evolutionary unfit behavior (and it is if we expect selection to explain everything) that gets exapted from other features, then this still does not explain how such behaviors emerge. That is, how can behavior emerge that transcends the imperatives of selection?

Moreover, who decides whether the burden of proof should be on selfish-gene theorists who deny the existence of genuine altruism or on those who theorize that genuine altruism exists? At best, selfish-gene theorists have not demonstrated that real altruism does not exist; rather they have only provided alternative explanations of the behavior in question. One problem here stems from the nature of our definitions. If we grant selfish selection all-encompassing explanatory status, then we must call extreme forms of altruism maladaptive or misappropriated. However, if natural selection stands as one of many explanatory means, then we can say these behaviors transcend current Darwinian explanations, and then we can look for other ways to explain how radical altruism might emerge.

4.6 Getting at the Root of Maladaptive Morality: Nonrealism

Before looking at the potential linguistic mechanisms that might help produce forms of radical altruism, we need to review some basic questions. What does it mean to categorize these behaviors as maladaptive or manipulated? To reiterate, it means natural selection does not directly produce the behavior in the altruist. Perhaps selection causes the manipulator to procure the behavior from the hapless altruist, or perhaps the altruist is simply out of sync with selection. Therefore, maladaptation refers to a behavior that is not fit or effective for the particular situation, function, or purpose for the altruist. Since natural selection concerns reproductive success, and radical altruism does not serve this, it stands as non-adaptive. This is especially evident when the altruism neither benefits the altruist nor her group in way.

However, this brief summary of maladaptive morality raises serious questions. First, why must we interpret *every* radically altruistic act as maladaptive from the perspective of natural selection? Some radical altruism may indeed be reduced to a maladaptation, where the behavior goes awry, and the altruist mistakes non-kin for kin. Nonetheless, it is still possible for an individual to sire numerous offspring, care for their fitness needs, and even the fitness of his grandchildren while still participating -- at least some of the time -- in many kinds of radically altruistic behaviors. In fact, this person's radical altruism may have no negative cumulative

affect on his reproductive fitness. For example, with some surplus of resources, he could anonymously serve in a soup kitchen once a week for 20 years, benefiting non-kin outcasts, while still fully providing for the fitness of his kin. Therefore, a simple answer might be that natural selection is not necessary to explain such behaviors all the time, and other mechanisms instead will help us understand them.

Thus, the "maladaptation" factor may not only bring us to the limits of natural selection, but it also highlights the claim that moral nonrealism or nihilism comes as a logical outcome of natural selection (Ruse, 2002; Sommers and Rosenberg, 2003). This moral nonrealism asserts that the moral beliefs, which might motivate radical altruists, are non-real or false. To quote Sommers and Rosenberg:

Nihilism consists in the following claims: a) normative terms-good, bad, right, duty, etc -- do not name real properties of events or things, either natural nor nonnatural ones; b) all claims about what is good in itself, or about categorical moral rights or duties, are either false or meaningless; c) the almost universal beliefs that there are such properties and that such claims are true can be "explained away" by appropriate scientific theory. (Sommers and Rosenberg, 2003: 655)

Though these authors advocate a "nice nihilism" (obviously without any good or real reason), and though their position would work if natural selection were a tenable theory of everything, (I claim it is not), there are serious problems with their position.

Generally speaking, scientists are realists, or critical realists. While recognizing that culture influences our knowledge, scientific realists believe they work with a real world and that their theories and mathematical formulae approximate the world as it really is. The success of the scientific method attests to the real nature of the world and our relatively accurate understanding of it. If scientists claim that the physical world is real but that moral ideas are non-real, they introduce a bifurcation and dualism into our understanding of the world, or at least into our philosophy of science. Conversely, a thoroughgoing scientific realist would tend to look at morality as something that is real. However, before we develop this point to a sufficient degree, we need to consider more fully what moral non-realism entails.

For issues of right, wrong, good, or bad, one problem comes from where one gets the added premise of moral evaluation. Moreover, acts of heroism, such as saving the drowning child or protecting the persecuted, attain universal appeal as the highest form of moral action. A simple but horrible moral story will highlight our problem. Say a man witnesses a child drowning in shallow water. He could easily save the child -- at minimal risk, but instead he waits till the child drowns, and then he commits indecent acts with her, and afterwards performs a cannibalistic ritual before leaving the remains of the victim -- a victim he could have saved. To make matters worse, the child's parents have driven to a nearby market get some extra food for their family picnic, and on their way back they get stuck in traffic. From their car, they witness everything, and finally while running to the waterfront to intervene, they fail to catch the perpetrator. They are left in a state of ultimate despair, not only having lost their beloved child, but also having witnessed her death and victimization at the hands of a man who could have saved her.

Clearly, moral non-realists would never condone such acts. This is not the problem. Rather, virtually all moral non-realists would strongly condemn such behavior; however, they would also have no *real* justification to categorize this behavior as *wrong*. They could *not* say that it is *true* that this behavior is *evil*. In fact, they would have to say that any normative claims that call such behavior as evil, bad, or wrong, are false or meaningless. Thus, a serious contradiction results from this thinking. Regarding behavior that we consider the epitome of wickedness, or for that matter, behavior that we consider to be the highest good, a moral non-realist has no real or true way to affix a moral evaluation to it. However, for a moral realist, the behavior described above is not simply repugnant; he also has good reason to believe the behavior is really wrong.

This discussion highlights the fact that affixing a real moral judgment to behavior stands as a central problem in evolutionary studies of altruism. For example, evolutionary philosopher Michael Ruse asserts that no real and factual foundation exists for the added moral premise. He says:

Once we see that our moral beliefs are simply an adaptation put in place by natural selection, in order to further our reproductive ends, that is an end to it.

Morality is no more than a collective illusion fobbed off on us by our genes for reproductive ends. (Ruse, 1991: 506)

Ruse is not claiming that the content of our morality is unimportant, but rather that even though we experience morality as psychologically and factually real, this experience is an adaptive deception. Because our reproductive fitness requires us to cooperate, our genetic make up guides us to act as if morality has a real foundation -- even though it does not. In this scenario, selection conferred reproductive fitness on the best cooperators, and we inherited their legacy. Ruse's clever move altogether avoids the *is-ought problem* raised by Hume; that is, Ruse neither tries to derive an *ought* from an *is*, nor does he attempt to justify any real *ought* at all. We think *oughts* are real because we believe and experience them as authentic at the deepest level, but they come not from any *is*. Hence, morals are adaptive fictions even though we treat them as if they were real.

Although Ruse's views are important and interesting, they are problematic also. As already mentioned, a first problem concerns what actually causes this altruistic behavior; another concerns how Ruse and others potentially conflate realism and non-realism in scientific thinking, and these two problems are related. Does a non-real adaptive fiction produce sacrificial and therefore apparent altruism, or is there a real moral cause for nevertheless rare but also real altruism? Midgley (1991) suggests that a society cannot exist, thrive, and overcome conflict without morally guided cooperation, and perhaps this provides a firmer foundation to add the moral premise. That is...

Morality needs, not just conflicts, but a willingness and a capacity to look for shared solutions to them. As much as language, it seems to be something that could only occur among naturally social beings. (Midgley, 1991: 12)

Hence, perhaps adding the moral premise comes not from a biological *is*, but from a higher level cultural *is*, that cultures cannot emerge and exist without rules of cooperation. Moreover, though we may not always derive these cultural rules from biological imperatives, we still articulate moral rules through language in ways that cohere with our biological needs. Thus, these rules often look after the psychobiological needs of individual members of a culture in ways that promote the fitness of the group as well. The explicit rule "obey your parents" not only

protects children from danger, but it also protects the physical health and social unity of the group. Moreover, this particular command only becomes problematic for SGT if our parents tell us to show altruism to strangers. In any case, the issue of how we justify such a moral premise impinges on the thesis of this chapter, but also transcends it. Therefore, before we discuss the linguistic and cultural mechanisms that might help produce radical altruism, we need to briefly consider a case for moral realism and its relationship to language.

4.7 The Relation Between Moral Realism and Human Language

In this section, I will test the following claim. Some scholars have failed to consistently apply critically realist and therefore fully scientific thinking in their study of natural selection, altruism, and morality. This failure reveals an even deeper reason why these scholars label the most highly esteemed and radical forms of altruistic behavior as only apparent, maladaptive, or manipulated. That is, such behavior is not only "selfish" altruism, but there is also no true, real, or factual moral reason for it because real morals do not exist. However, a more full-bodied critical realism may make it possible to get rid of this bifurcated thinking regarding morality and reality. To quote the seminal critical realist philosopher, Roy Bhaskar:

Ideas, and ideational connections (including category mistakes, logical contradictions etc.), are a part of everything, and everything is real. To deny the reality of a part of everything (of anything), such as ideas (or say persons, or consciousness, or agency, or values -- or mind, or body) extrudes or detotalizes it or them from the world, that is the rest of the world of which they are in principle causally explicable and causally efficacious parts. This inevitably produces an implicitly dualistic or *split* ontology. (Bhaskar, 1997a: 139)

In this basic claim for the reality of ideas, Bhaskar does not say that all ideas "possess the same ontological type, or categorical status" (1997a: 139). He denies this. However, he claims that though we may not be able to distinguish the reality of moral ideas (categories) directly, we can experience them through the efficacy of their effects, or as I will argue below, through real categories that impinge upon us in a moral way.

For Bhaskar, a reason we deny the reality of moral ideas comes from what he calls the epistemic fallacy -- the conflation of ontology and epistemology.

Empirical realism is underpinned by a metaphysical dogma, which I call the epistemic fallacy, that statements about being can always be transposed into statements about knowledge of being. (Bhaskar, 1997b: 16)

For this discussion, the question is: by saying that there is no real basis for morality, do moral non-realists commit the epistemic fallacy, confusing the ontology of moral ideas with the epistemology of moral ideas? Clearly, according to Bhaskar's thought, denying real status to morality implicitly expresses the epistemic fallacy because it produces a dualistic ontology. For example, Ruse (2002) splits the nature of being; he says that substantive morality is ontologically not real, but that we mistakenly or knowingly view it nevertheless as epistemologically real. Incidentally, that Ruse and others commit this fallacy still does not show the real ontological status of moral ideas or categories. Thus, in the following discussion I will expand on the nature of the moral non-realist's epistemic fallacy and briefly suggest potential pathways that may lead us towards a basis for the real ontological status of moral ideas.

Indeed it is important to show that moral categories are real in order to refute moral non-realism, but it is also essential to show how the thinking of Ruse and others reveals the epistemic fallacy on this subject. Ruse's denial of moral reality expresses the epistemic fallacy because he conflates "what is" with "what is known," or "what we can and cannot know." This fallacy comes from the epistemological limit imposed by a strict or inflexible application of Hume's Law -- that says we cannot move from *the way things are* to *the way things ought to be*. That is, we cannot derive an *ought* from an *is*. Moreover, Ruse has made it explicitly clear that he is attempting to follow Hume in his moral non-realism (2002: 162-167).

Interestingly, the epistemic principle that says we cannot know an *ought* from an *is* preempts the very possibility of moral ideas attaining real ontological status. However, this epistemic premise does not actually show that moral ideas are non-real. That is, the *is-ought rule* is not an ontological statement about what exists, but

rather an epistemic statement about what we can know -- a rule about what can and cannot be known. However, if we follow an *a priori* rule that says we cannot know or derive moral reality from biological, social, or psychological reality, then we dangerously confuse our epistemology with what may or may not be ontologically real. In fact, rigid use of the *is-ought rule* treats an epistemic premise as if it were ontologically real. Thus, we are doomed to never derive an *ought* from an *is* because we act as if the *is-ought rule* is ontologically real when it is actually an epistemological premise. We would do better to find *a posteriori* ways to approach the question of moral reality instead of answering the question in advance with an epistemic preemption. Therefore, this kind of moral non-realism clearly exhibits the epistemic fallacy, the conflation of ontology and epistemology, as described by Bhaskar (1997b).

This state of affairs raises two questions. First, is this epistemic rule right or not? That is, could it be a principle derived *a posteriori* from the real world (instead of just being a presupposition), and second, if it is not right, how might we ascertain the ontological real status of moral ideas? Though a full discussion of these issues exceeds the limits of this chapter, a simple problem for the moral non-realist may be that he only ascribes real status to that which he can perceive directly. However, moral ideas may be perceived indirectly, through their actual or potential effects. If this is the case, we should use a causal criterion (not a perceptual criterion) when judging the ontological status of moral ideas (Bhaskar, 1997a). Moreover, there may be a set of categories in the real world that might cause some kind of general and real morality to impinge upon us.

Again, in this thesis, we cannot fully delve into moral realism and its relation to language, but we can briefly hint at the role language might have in a morally real world. First, we must note that even the hard facts that moral non-realists consider to be real are nevertheless value-laden. A moral non-realist would likely say that factual discourse cannot be about values and cannot lead to values or morality. However, the critical realist acknowledges that facts stand as social constructs, not in the post-modern sense, but in the sense that culture and society influence our

epistemology so that values partly mediate our approximate understanding of real facts.

Moreover, even moral non-realists must make value-oriented judgments that also have a real sense about them. Bhaskar suggests (1987) the *is-ought rule* expresses a commitment to truth, and that when we expose a source of untruth, we evaluate it negatively, and move to eliminate it. For example, we can state the rule like this. "It is true that you cannot derive an *ought* from an *is* so any morals allegedly derived from facts are not real, and we thus *ought* to deny their real status." This can lead to circular thinking: the *is-ought rule* *is*; thus, we *ought* not derive any *ought* from it, except that we ought not derive any *ought* from any *is*, that is except from the ontological real status of the *is-ought rule* itself. At this point, we need to note that a softer version of the *is-ought rule* can help us to make cautious fact/value distinctions, instead of sticking us in a harsh fact/value dichotomy. In short then, the commitment to truth implied by the *is-ought rule* obviously leads to some prescriptions or values about actions.

In addition to the value laden aspects of the *is-ought rule*, critical realism also considers reality as multi-layered, multi-leveled, or stratified; hence, our methodology will differ depending on the ontology of the stratum in question. To elaborate, here is an example of stratification in rather simple terms: the chemical strata underlie the biological strata; the biological strata underlie the psychological strata, and psychological strata underlie the social strata. Varying methodology depending on stratum is implied in the multi-level selection theory of Sober and Wilson (1998) and in some memetic proposals, such as Dawkins (1976). However, these approaches may still be somewhat overly reductive regarding the question of a putative moral reality in that they keep selection central to their methodology. That is, the ontology of the object of study may require additional methods beyond Darwinian selection, especially since culture and morality may have partially broken free from selection's leash.

"The essential point is that each stratum must be investigated by a methodology which is determined *a posteriori* by its ontology" (McGrath, 2002: 224). Hence, though using a selectionist approach may reveal insights about all the different

levels, we should use this approach not only because it works for biology but also because the ontology of psycho-socio-moral reality suggests that we use it. Biology influences psychology and sociality; however, there are still aspects of psychology, sociality, language, and morality that may transcend biology and the limits of a selectionist approach. In these cases, we may need a different methodology or a combination of methodologies to match the appropriate stratum of study.

In sum, though we are far from beginning to establish a case for moral realism and its relation to language, we can perhaps say that if moral reality exists, it might consist of the following interlocking factors. First, the natural world exists in a real way, (a way that is also stratified), and science brings us an approximate but also real understanding of the way the world *is*. Second, as mentioned before, even the *is-ought rule* implies a kind of moral obligation to discern what is true or right. Third, moral 'beauty, goodness, or truth' may exist in this world in ways that we can actually discover them, rather than just invent them. That is, we may find *a posteriori* ways to infer 'goodness or justice' from the ordering or patterning of the natural world (McGrath, 2002). Fourth, language will play an essential role helping us to discover, describe, and abide by any kind of moral reality. Fifth, the Golden Rule might express one aspect of moral reality, with a biological substrate that expresses visceral empathy and a psychological substrate that expresses cognitive empathy, which is enhanced by language. There will also be a social substrate of empathy, codified in a shared language, and expressed in interdependent human relationships and describable and extended by the Golden Rule.

This moral realism may still be adaptive (Harms, 2000), and it will partly mesh with kin selection and reciprocal altruism. However, because culture in language has the potential to release us (to a degree) from selection pressures, and since cultural change can outpace evolutionary change, moral reality -- partly expressed by a *g code that ratchets up as a response to that reality* -- may also explain how we can transcend the influence of selfish genes and groups. Moral realism may also give us reasons *why not* to follow ethics similar to the Third Reich, which exalted fitness, intelligence, health, and genetic purity as the highest standards of morality. See Weikart (2004).

In addition to these factors, we can look at moral reality with causal criteria in mind. Of essential interest here concerns how we linguistically categorize and approximate reality in ways that may causally influence moral behavior. In Bhaskar's critical realism, it is important to note that even inaccurate categorizations are still real in that they are part of the whole and real world. Moreover such inaccurate or accurate categorizations will also affect our behavior, and in that sense they will have an efficaciously and causally real influence on our morality.

How do our categorizations affect our behavior? If we use language to categorize an ethnic group of humans as animals, putting them in zoos, buying and selling them, or using them as slaves or food, then this mistaken categorization clearly has moral implications. Moreover, humans would definitely use (and have used) language to make such false categorizations, and we would use language to make the accurate categorizations as well. We can respond to these language-based categorizations in various ways, but in general if we categorize a group of people as animals, then we will most likely treat them as animals as well. Thus, the way people linguistically categorize morally latent situations -- whether accurately or inaccurately -- will strongly influence how they behave. Slave owners categorized slaves as subhuman, or non-human property and treated them as such. Moreover, from a realist perspective, the if moral psychosocial causes are real, and depending on the accuracy of the categorization, the behavioral results may be more or less in line with the actual categories in the real world. All in all, if moral reality exists, language can enable us to accurately understand and act in accordance with it.

4.8 The Role of Language in Developing Altruistic Personality

Above I have briefly suggested how critical moral realism might relate to how we use language to categorize the world. This is an attempt to move us beyond the contradictory and non-scientific non-realism or anti-realism of some scholars. Such a critical moral realism does not need to be naïve, and it holds promise for future development; see: (Millikan, 1984; Bhaskar, 1997b; Norman, 1997; Harms, 2000).

On the basis of previous discussion, we may now more easily move ahead to discuss how language might influence real altruism in its more radical forms. Current research on altruism and morality strongly suggests that selfish gene and group selection theory provide only a partial and therefore incomplete theoretical and empirical means for understanding human altruistic and moral behavior. For example, see: (Baston, 2002; Kagan, 2002; Monroe, 2002; Oliner, 2002; Post et al., 2002; Wyschogrod, 2002). This research shows that human altruism is complex and requires insights from numerous disciplines, such as psychology, sociology, neurology, and as this thesis argues, from linguistics as well. Moreover, sociobiology, despite its success in explicating altruistic behavior through inclusive fitness and reciprocity appears to sketch only a partial picture of this complicated problem. Thus, similar to studies on the origin and evolution of language, which encompass many disciplines, we may require a multi-disciplinary approach to understand the evolution and grounding of altruism as well.

Regarding linguistics, I have *not* contended in this thesis that language supplies the only mechanism for facilitating and maintaining altruism in human individuals and groups. However, I have asserted that language is one essential in a matrix of factors that makes human altruism and morality possible. Moreover, my approach contends that while non-human species may exhibit altruism and substrate morality (de Waal, 1996; Bekoff, 2001), human language has the power to greatly extend human altruism and morality. To avoid tendentiousness, we must note that language also increases our power to deceive and manipulate others, but does not negate the thesis here; it only makes it more complicated.

A simplistic approach might suggest that people have linguistically based moral codes, and when faced with a moral dilemma, they access the code as it relates to the dilemma, and try to act accordingly. This straightforward approach -- *access and act upon the code* -- may describe the way altruists behave in some situations. However, Monroe (2002: 110) claims that research on the "altruistic perspective" shows that the process for producing genuine altruism or *other regard* is rather different. By way of background, Monroe has done extensive research on holocaust rescuers, and she exactly defines what she means by altruism. First, altruism

involves action; hence, altruistic thoughts and intentions do not constitute genuine other regard. Second, the action must have either a deliberate aim to benefit the other, not the altruist. Third, the result of the altruistic act is more important than the intent of it. Fourth, the altruism is unconditional without expectation of reward, and fifth, the altruist's risk of loss of well being always accompanies the other regarding behavior. Regarding SGT and GS perspectives, Monroe states the following:

Reciprocity, clusters of altruists, and group or kin selection -- all typical explanations from evolutionary biology and economics -- have no predictive power for altruists, nor does the cost benefit calculus or the psychic gratification hypothesis of economics... These traditional explanations, then, offer limited explanations for altruistic-like acts by rational actors but fail to explain altruism itself. (Monroe, 2002: 108)

Thus, Monroe claims that a genuinely altruistic personality emerges from the altruist's "cognitive-perceptual frameworks"(2002: 109), especially regarding the altruist's self-perception in relation to others. In this view, the risky altruism of holocaust rescuers does not come out of an agonistic choice between the actor's selfish desires and the demands of an other-regarding ethic. Moreover, Monroe claims this altruism does not come from a rational morality that attempts to overcome fight, flee, feed, or reproductive passions. Instead, the altruists' "decisions to risk their lives to help another appeared spontaneous" (2002: 109). This spontaneous altruist does not have time to calculate fitness costs, and this coincides with how a person responds in an emergency situation, like the aforementioned airplane crash where the rescuer only has time to act.

Thus, Monroe suggests that the self-perception or identity of the altruists constrained their choice to behave in radically altruistic ways. For example, the subsequent phrases depict the thinking of those who acted altruistically. "What else could I do?" "They were human beings just like you and me." Monroe quotes one rescuer's thoughts, which especially exemplify this mindset: "The hand of compassion was faster than the calculus of reason." These typical responses from rescuers caused Monroe to conclude: "identity trumps or at least severely constrains choice" (2002: 109), and this mindset seemed to be indicative of the bystanders (who watched and did nothing) and the Nazis as well.

Consequently, when confronted with a victim or sufferer, altruists constrained by their self-perception, spontaneously attempted to help a person in need. Thus, Monroe states that:

It is the altruistic perspective that most successfully explains altruism. It is the only factor that consistently and systematically predicts altruism among the individuals I interviewed. This altruistic perspective consists of a common perception of themselves as individuals strongly linked to others through a shared humanity. This perspective constitutes such a central core of their identities that it leaves them with no choice in their behavior toward others in regard to ethical issues. (Monroe, 2002: 110)

Regarding language, the "altruistic perspective" does not indicate that language plays no role in making altruism. The first chapter of this thesis outlines the role that language plays in moral/altruistic socialization (and identity development), and although Monroe has not found socialization dispositive, she has still not written it off as insignificant. In fact, in her final analysis, she gives linguistic ability a significant, if not central role, in the development of the altruistic perspective as it relates to identity.

Regarding socialization, Monroe mentions the need for more research regarding how socialization helps develop the altruistic perspective, and she suggests focusing on how we create our social identity and categories of self. As for actually developing altruism she says:

I suggest we stress socialization that emphasizes an identity based on common membership in a shared humanity rather than an identity that lauds group differences. In policy terms, this would privilege humanistic socialization over multiculturalism. (Monroe, 2002: 112)

Of course, this relates to our concern with the Golden Rule and the factor of virtual relatedness. In fact, one of the Holocaust rescuers whom Monroe cites says: "They were human beings just like you and me" (2002: 109). This Golden-rule type of categorization, which linguistic argument structure enables and extends, correlates with the thesis that language plays a central role in developing and maintaining altruism and morality.

Moreover, though Monroe asserts that identity constrains choice, this does not rule out the possibility that an altruist can access and act upon a moral code in situations that force choices between altruism and selfishness. Real altruism most likely emerges from a complex process. Thus, a potential altruist can access and act on the moral code spontaneously; moreover, his behavior can become automatized -- like the way native speakers use their first language. The altruist could express his altruism and identity simultaneously through linguistically dependent thoughts, such as: "That could be me, and if I don't help him, who will?" Moreover, the altruist may respond spontaneously because of many factors such as a linguistically driven socialization process that emphasized shared humanity. Additionally, he may have experienced various forms of altruism with kin and non-kin, and his caregivers may have constantly taught him a moral code that emphasized positive altruism to outsiders. In a word, a person may become a real altruist through a complex process -- involving many interdependent factors -- with language at the core.

In addition, not all radically altruistic choices need to be spontaneous. Responses might vary between spontaneous and deliberative depending on the kind of choice and situation presented to individuals. Even if the initial response to rescue a victim is constrained by identity, the moral code can still help create that identity. Moreover, some altruistic responses are prolonged and may require constant attention, planning, and reaffirmation. For example, the Ten Boom name is well known in Holland for subversive resistance during the Nazi occupation, and their home was called The Hiding Place -- for fugitives hunted by the Nazis, including members of the Dutch underground and persecuted Jews. Though it is hard to ascertain altruistic motivation, which differs between altruists, members of the Ten Boom family apparently found motivation for their altruism in their strong ties to the culture of the Dutch Reformed Church. Moreover, members of the family were well schooled; they studied anti-Semitism and worked to improve Jewish-Christian relations in Holland.

The point here is that this family expressed altruism as a way of life, informed by linguistically dependent education and clear and constant religious instruction.

Obviously, the content of their spirituality influenced their way of life, and Monroe (2002: 113) emphasizes the importance of the *content* of religious values. "It is the *content* of the spiritual life that is critical, not merely having a spiritual life" (emphasis in the original). Thus, it seems sufficiently clear that the Ten Boom family expressed their worldview through their altruism. Moreover, though the family was eventually imprisoned for what they did, it appears that they and their friends were able to rescue approximately 800 Jews and numerous underground workers through their "safe house" efforts.

We must note that this work exacted a constant cost on all the altruists and an ultimate cost on some of them, and this raises the following question. How does one keep up the courage to sustain prolonged altruistic efforts like this? The answers will vary depending on the altruist; however, the Ten Boom family undoubtedly got courage to continue their altruism from the content of their linguistically dependent religious faith. For instance, after the family was imprisoned in February 1944, the father (84 year-old Casper Ten Boom) was asked if he knew he could die for helping Jews. He replied: "It would be an honor to give my life for God's ancient people" (Corrie Ten Boom, 2006). Moreover, before the family and collaborators (a total of 35 people) were moved to separate cells, they were all together in a large room, and Casper read Psalm 91 to the group. (Psalm 91 gives encouragement to those experiencing severely stressful situations.) This example shows how an altruist may access his code to sustain, justify, or affirm his actions through language.

Later on, Casper's daughters Betsie and Corrie were sent to the Ravensbrueck concentration camp in September 1944; Betsie died there in December, and Corrie was released soon after that. The main point here concerns how the sisters maintained altruism in the concentration camp, for in spite of atrocious conditions, the two were able to inspire and lead fellow prisoners. Specifically, it is noteworthy how these two maintained courage and altruism in the face of adversity. For example, Corrie was able to sneak scriptures into the camp, as the guards accidentally skipped performing a thorough search on her when she entered, and the sisters used the book to aid their emotional and psychological survival. That is,

they accessed a code with language to constantly maintain and affirm their altruistic perspective.

Moreover, when Corrie was released she apparently continued her altruistic behavior by advocating altruism, reconciliation, and forgiveness towards those who had been her captors (Boom and Buckingham, 1986). In fact, at a meeting after the war, Corrie saw one of the guards who had beaten her sister at Ravensbrueck, and after the meeting he approached her and asked forgiveness for his cruelty. It is important to note Corrie's response, which was not spontaneous, and though Monroe's ideas of identity and the "altruistic perspective" may be relevant, it appears that Corrie Ten Boom needed to mentally and verbally access her moral code in this case. Her response was *not* "What else could I do -- this person was just like me while at Ravensbrueck?" She mentions deliberately accessing her moral code, and then at first she mechanically and woodenly grasped the guard's hand, and as their hands touched she began to weep and then spoke words of forgiveness to her former captor.

This anecdote of Corrie's deliberate altruism highlights a vital issue in altruism studies. Altruistic reconciliation between severely estranged parties, though uncommon and difficult, nevertheless seems possible. Moreover, though further exploration of this idea takes us beyond the limits of this chapter, it appears that when humans reconcile in radically altruistic ways, we may not simply empathically categorize and identify with our interlocutor. Perhaps in these cases, even some who have developed a highly altruistic perspective will need to calculatively access a linguistically based moral code in order to forgive in this way. In any case, this example raises an interesting and vital issue for studies concerning language, altruism, and morality, highlighting situations where we may need to access a linguistically dependent moral code.

In addition, the Ten Boom family is illustrative here for many reasons. First, they demonstrate that some people learn and access a linguistically dependent moral code to develop and sustain an altruistic personality. Thus, for the Ten Boom case, we cannot easily deny the essential role of a linguistically dependent moral code for developing an altruistic perspective. Second, the above historical events

demonstrate the ability of the moral code to sustain and affirm the behavior and attitudes of the altruists. For example, altruism that saves 800 people over a long period of time under severe stress and in great danger requires sustained physical effort and emotional energy. Thus, when an altruist constantly accesses her code, its exhortations, encouragements or inspirations, she significantly sustains and affirms her identity and her behavior. Thirdly, the above example stems from a religiously based code, but we can generalize the principle here. Hence, religious scriptures, secular proverbial wisdom, the verbal advice of friends and family, an altruistic philosophy of life, and linguistically encoded (and oft recalled) habits of altruistic thought can develop, sustain, and affirm altruistic behavior. Thus, a theory of language and morality where an altruist accesses and acts upon a moral code strongly complements Monroe's view on the "altruistic perspective."

4.9 The Altruistic Perspective and Linguistic Ability

Along these lines, Monroe's research suggests an additional way that linguistic ability helps develop the altruistic perspective. In a word, she discusses how "language categorization leads to the development of ethics" (2002: 120). Her view differs from but still coheres with the linguistic and categorical moral realism discussed above, and we can summarize her view in the following five points. First, based on innate dispositions, humans need self-esteem, and we fulfill this need when we behave consistently and thus form our identity through acts that confirm our sense of self. Second, we desire decent treatment from others, and third, the universal human abilities of cognitive categorization and linguistic communication enable us to comprehend this desire in others and ourselves. Fourth, this forms a coherent basis for treating others, as we desire to be treated. Fifth, this reciprocal ethic forms a fundamental element of the human ability to communicate and to distinguish boundaries through categorization (as opposed to some rational or religious sense of duty). Though there are different points of departure, this basic outline of Monroe's theory overlaps nicely with the previous section describing moral realism, categorization, and language.

Additionally, Monroe's work supports this thesis because she not only claims that humans can show radical altruism to non-kin out-group members, (a central claim of this chapter), but also altruists significantly employ language in this process. At one level, regarding the altruistic perspective, two key issues emerge relating to how psychological, social, and religious factors affect the development of the altruistic perspective. Concerning psychosocial empathy, Monroe claims that empathy itself is not enough, for a twisted mind can use it to maximize suffering. Rather, altruists need to act on empathy as if they were showing it to themselves. Moreover, religious affiliation does not necessarily foster the altruistic perspective, for humans can arrest even the most radical altruistic teaching into the service of the in-group at the expense of the out-group. Thus, as mentioned above, Monroe claims the *content* of the religion must also encourage the altruist "to place the ultimate value on human life before it will trigger the altruistic perspective" (2002: 113).

Here we see how Monroe theorizes that language can play an absolutely critical role in the development of the altruistic perspective. The altruist comprehends himself as standing in the subject category "I," and he sees the recipient of the altruism as standing in the object category "You." Moreover, the altruist links these two categories as they share the common category label of "human." In addition, if the altruist cognitively switches these categories, he can activate genuine empathy in the categorical terms "You could be me." Thus, the altruist may see pain in another and desire to relieve it because he thinks, "You could be me" or because he actually feels that suffering. If a person has been raised up in an altruistic perspective, his "I" category will be intrinsically linked to the "you" category. If the "you" category extends to all humans, then maltreatment of anyone in the "you" category will place pressure on his altruistic perspective, especially when he stands in a position where he can help a sufferer. That is, the categorical linking of "you" and "I" will place pressure on the altruist's sense of self-esteem, and he will want to act congruently with his identity. Hence, I see you as me, and cannot bear to allow your suffering, so I do to you as I would be done to.

This echoes Buber's (1971) philosophy of personal dialog, characterized by two attitudes that one can have about relationships. *I-It* relations are subject-to-object relations, where we would treat others selfishly as objects, and where we are separated and detached from others. *I-Thou* relations are subject-to-subject relations where we affirm the shared humanity of others. Though there are limitations to Buber's work, it does succinctly express key ideas discussed here, and it also coheres with the idea that linguistic ability is essential for developing real altruism.

The problem with these approaches concerns how to test them against evidence, and there are two potential ways: one negative and one positive. First, on the negative side, we can see how racists have used language in propaganda, especially where they have tragically dehumanized others. Clearly, it is valid to ask how racists have recategorized people in order to treat them differently or cruelly. For example, Nazi Germany used a linguistically dependent propaganda machine to dehumanize Jews, recategorizing them as "untermenschen" or subhuman; thus, Jews were demonized and considered polluters of racial bloodlines. The US military has used language -- in the form of racial epithets -- to dehumanize their enemies. Therefore, this shows rather straightforwardly how humans can fully exploit language to promote anti-altruism. Language is sharply double-edged in this regard. The work of Staub (1992) and Glass (1999) demonstrates that when we falsely recategorize people, we create a culture of inequality, persecution, and anti-altruism where fellow humans become "unworthy of life."

Secondly, on the positive side, we can also research how humans use language to promote empathy and the kind of categorization that harmonizes with the Golden Rule. For example, we can show how aid organizations use the language of empathy and common categorization to promote their altruistic programs. Food for the Hungry's web site describes the pain of a hungry child, eliciting empathy, and encouraging readers to donate and sponsor a child. UNESCO aims to bring about positive dialog between nations based on shared values and the dignity of every civilization and culture, a dialog that would have out-groups positively categorize each other in ways that cohere with the Golden Rule. Since these groups are

basically marketing altruism, we may find some ways to evaluate their effectiveness. Moreover, we can evaluate how real altruists verbally categorize their beneficiaries in order to see if how they categorize others correlates with their altruistic behavior.

Therefore, regarding linguistic means for promoting radical altruism, we may find relatively clear ways to show in detail how language helps us ratchet up empathy, or at least how we can use language to promote altruism. Sorokin (2002) called this "extensivity" where we can conjugate and extend the boundaries of "self" and "other" in many ways that would be impossible without the abstract and displacement characteristics of language. Thus, we can extend the "shared humanity" category to every person on earth -- even those who deny it or flout it -- thus enabling us to moralize about the ethical treatment of criminals or even tyrannical dictators. This is similar to what Corrie Ten Boom did by explicitly promoting forgiveness and reconciliation to those who imprisoned and persecuted her. Moreover, language enables us to extend the category of "shared humanity" to creatures with "shared intelligence" or "shared life." For example, the World Wildlife Fund's web site refers to the panda as lovable and charismatic. This is anthropomorphizing (a form of re-categorization), and another example of how language enables us to expand or extend categories. Thus, extensivity stands as one key factor regarding the potential of human altruism and morality. In fact, Sorokin claimed the lack of extensivity as expressed in in-group altruism at the expense of the out-group, (group selection if you will) has caused more anti-altruism than *any other trait* in all history.

The point is that if we lack the trait of extensivity, we may not afford others the category of equal or fully human. This at least opens the door so that *the unequal* can be bought and sold as slaves, persecuted and tortured as subhuman, killed and maimed as evil beasts, or incinerated as gene pool polluters. This may be an oversimplification, but as just mentioned, there seem to be tangible ways to show that non-extensivity serves as a common denominator in substantial movements of non-altruistic behavior. Language enables us to label reality as fantasy, and vice versa. Thus, we can easily categorize those who are superficially different from us

as "outsiders," and with language the "you" can also become the "un-you" who deserves *not* the kind of treatment that I would afford myself. Hence, we can rearrange and realign the categories to fit our malicious intentions.

However, such values and categories, though contrary to any sense of universal human rights, are still values -- that is, they are valued by some. This does not make them right, but the interesting issue, from a descriptive perspective, is that these values (or un-values) also seem to find their strength *in and through* the human linguistic ability. Thus, studies in language categorization may lead us develop theories of morality and ethics. Moreover, research on linguistically dependent altruism appears to show serious potential for further work on the origin, grounding, development, and maintenance, of altruism and morality in human beings.

Finally, though the universal age of reason has emerged as a failed idea of the Enlightenment Project because cultural perspectives partly relativize our understanding, we still have good reason to consider morality in light of critical realism. In addition, in view of what has been stated above regarding morality, realism, language, and categorization, the common features of human languages may provide a partial basis for an elemental or universal morality. This does not need to be a naïve or presumptive claim, but it can emerge *a posteriori*, as we approximately describe real biological, psychological, and social categories common to all humanity. Moreover, as humans universally and linguistically categorize others in altruistic (and egoistic) and moral (and immoral) terms, it seems fitting that in the next chapter we consider universal features of language and how they may influence human altruistic and moral ability.

CHAPTER 5

THE PROPERTIES OF LANGUAGE THAT EXTEND MORALITY

5.1 Linguistic Properties of Language that Enable or Enhance Morality

This thesis focuses on language and morality related to evolution, such as how language influences kin-oriented altruism, reciprocal altruism, and group selection, or how language helps facilitate altruistic behavior that transcends selfish genes or groups. Morality remains a central problem in evolutionary theory; moreover, basic questions about the origin of language remain challenging and unanswered. Since morality depends on language, the origin of language and morality might relate to each other, and thus linguistic properties that enable or extend human morality may also relate to our moral origins. For example, *if* properties in predicate-argument structure enable or extend human moral action beyond kin, then one can argue that the origin of morality beyond kin depends considerably on this property. If any moral ability depends on linguistic properties, we may gain novel insight into our moral origins and abilities by exploring how specific properties of language enable specific moral abilities. Hence, this chapter describes and expounds upon how numerous linguistic properties enable and support altruistic or moral thinking and behaving.

5.2 Similarities and Differences Between UG and Universal Morality?

To begin, we can inquire into how Chomsky's (1965) Universal Grammar (UG) relates to morality. This is a complex question, perhaps demanding a whole

separate discussion; hence, this section on UG will be brief. First, we must clarify what UG actually means. Jackendoff (2002: 70) calls UG the initial state of a language learner, "as the aspect of the human mind that *causes* languages to have the features in common that they do." Thus, Jackendoff interprets UG as the innate brain structure that prespecifies "the form of possible human grammars." This neurological prespecification also allows humans to learn languages; and hence, "the grammar-*acquiring* capacity is what Chomsky claims is innate" (2002: 72). With this clarification, we can see why Jackendoff asserts that the term UG may mislead us, and he thus suggests the phrases "seeds of grammar" or "metagrammar." For this discussion, we must focus on how the *innate grammar acquiring capacity* might support human morality.

Hence, with these ideas in mind, we should state our questions about UG and morality in the following ways:

- Does the innate grammar acquiring capacity help prespecify the form of possible human moralities due to the way it prespecifies the form of possible human grammars?
- Is there a prespecified mental metamorality -- separate from UG -- that limits the form of possible human moralities and provides humans with an environmentally triggered innately guided morality-acquiring capacity?
- Do innately determined characteristics of the human mind cause human moralities to share universal features?
- Is the human brain pre-wired with the seeds of morality, separate from general cognitive learning capacities, that in varying degrees prespecify the structure of morality and/or provide learning strategies specific to the acquisition of morality?

These are most challenging questions that we cannot currently answer very well.

Nevertheless, we can begin to answer them by focusing on what empirical evidence actually supports UG, recapping previous discussion in light of these questions, and comparing how UG and language acquisition parallel morality acquisition.

5.2.1 Empirical Support for UG

Regarding empirical support for UG, no simple fact stands out that completely supports the theory; however, as Jackendoff (2002: 101) claims, a number of pointers taken *en masse* offer "an overwhelming case for some degree of biological

specialization for language learning in humans." He asserts that if researchers acknowledge the real complexity of language, and if they seek to understand what natively predetermined tools we use to learn and process language, then we can come a long way to more fully answering specific questions about UG. To begin answering these questions, Jackendoff outlines four areas of research that give linguists solid reasons to push forward with the UG project. They are (1) species-specificity, (2) the characteristic timing of acquisition, (3) dissociations, and (4) language creation.

First, species-specificity underscores why evolution of language studies would be much less significant -- if humans were not the only creatures with language. However, language is unique to us, especially fully syntactical and combinatorial language. For example, Hauser, Chomsky, and Fitch (2002) claim that humans uniquely apply general recursive ability to language, and if researchers compare recursion in other species (navigation or social relations), this comparative approach may give us insight into the origin of language. Jackendoff maintains that species-specificity, the lack of recursive language in even the most intensively taught primates, and the existence of language even in small-brained humans or those who have undergone hemispherectomy, demonstrates basic evidence for innate hard-wiring for language in humans. In short, the uniqueness of human language stands as one beam of support for universal grammar.

Besides species-specificity, Jackendoff discusses a second strand of evidence that supports UG -- stemming from the characteristic timing of acquisition. All normal children acquire language, but adults vary greatly in their ability to learn secondary languages (or even first languages in the cases of feral children). Moreover, regarding evidence for Lenneberg's (1967) critical period, Newport states that

... fluency in certain aspects of second language acquisition correlates inversely with the age at which one starts to learn. On average, people who start at age six get better than people who start at age twelve (Newport, 1997, cited in Jackendoff, 2002: 96)

This strongly suggests that something like UG emerges from a maturationally constrained, innately driven language prespecification in humans.

A third area confirming UG comes from dissociations -- how language disassociates from general intelligence, and how grammatical capacity maintains the same neural locus of control but can disassociate between modality of expression (in sign or speech) (Jackendoff, 2002: 97). For example, types of aphasia localize in the brain in both spoken and signed languages. Jackendoff explains, however, that one might expect "any cognitive activity" to localize in the brain, such as guitar playing or shooting a free throw in basketball. Nevertheless, even though sign uses aspects of the visual system, it also appears to use the same specific and localized cognitive resources as spoken language. Thus, aphasia affects language and the same part of the brain irrespective of the modality through which we learn it, and this suggests that part of the brain is biologically specialized for language.

Additionally, besides "dissociation between grammatical capacity and modality of expression," there are other "dissociations between aspects of language and general intelligence" (2002: 98). For example, linguistic savants are deeply retarded in other respects but have extraordinary talent for language learning, which implies dissociation between their general intelligence and language ability. In short, we should expect this kind of dissociation if grammar has an innate aspect to it.

Lastly, Jackendoff cites language creation -- as the most striking evidence that supports UG, and he gives three examples. First, a deaf child who is isolated from a signing community will tend to develop a communication system called Home Sign. Home Sign represents an improvised gestural communication system that expresses structure of "consistent word order and incipient morphological marking" (Jackendoff, 2002: 99). Jackendoff asks where this structure would emerge, if not from innate expectations about what language should look like. A second example of language creation comes from the creolization of pidgins. Bickerton's (1981) research indicates that the children of pidgin speakers use the poorer pidgin system as the basis to develop a much richer Creole. However, since all Creoles tend to employ grammatical devices not traceable to the pidgin, Bickerton argues that the children's innate grammatical expectations increase the complexity and novelty of the Creole grammar.

Finally, the famous case of Nicaraguan Sign Language exemplifies language creation -- where a dispersed deaf community of Home Signers came together through a school for the deaf. Over a period of 10 years a Creole sign language emerged from the rudimentary Home Sign pidgin, and it appears the metagrammar expectations of the signing community created this creole whose richness and complexity far transcends the Home Sign pidgin. In short, taking these four areas of evidence together, something more than general cognitive processes seems to be necessary for humans to learn languages, and as we will see below, this evidence will have some significant implications for human morality.

5.3 Evidence for UG Correlated to the Human Moral Capacity

With this brief background for the evidence for universal grammar, we can now look at how UG relates to our moral capacity. First, is morality species specific? Though many researchers agree that moral substrates or proto-moralities exist in non-human species (de Waal, 1996; Bekoff, 2002; Preston et al., 2002), scholars also agree that morality is a unique human trait (Harms, 2000; Katz, 2000; Preston, 2002; Goodenough and Deacon, 2003). In fact, our capacity for morality might stem from our ability to convey it with the complexity that language avails. Moreover, other species do not employ an elaborate moral code like humans, which we elaborate by our ability to name an infinite number of things as "good" or "bad," a point that we will discuss below under the rubric of language creation (and morality creation). In a word then, in as much as morality depends on the uniquely human aspects of language, those aspects of morality must be species specific also.

Second, is there a characteristic timing of acquisition for morality similar to what appears to be the case for language? This question has been answered to a rather significant degree in Section 2.3 of this thesis, which focuses on a parallel between moral and linguistic ontogeny in humans. As that section outlined, though future research may hold promise concerning the parallel between linguistic and moral development, clear evidence already exists that moral development is maturationally constrained (Schoen, 1994). This is a rather common sense point of

view, and cultures generally deal with moral issues maturationally, treating at least some moral offenses differently, depending on the age of the perpetrators.

Moreover, as moral issues become more complex, it is obvious that one must have the grammar and vocabulary to understand and discuss them. Generally therefore, grammar and morality parallel each other regarding the timing of acquisition.

Thirdly, regarding dissociation, and reiterating points made in Chapter 2, Anderson (1999) cites evidence that patients with adult-acquired damage to their orbitofrontal cortex (OFC) could reason well about hypothetical moral dilemmas. However, they failed to use this moral knowledge in real life, but subjects that endured early-onset damage made excellent recoveries, but later began to display severe moral dysfunction in real-life and moral reasoning. This evidence for maturational and localized constraints on moral development and ability gives significant credence to the idea that morality partly parallels UG. Moreover, though we cannot dissociate morality into different modalities of expression like sign and spoken language, people with adult-acquired damage to the OFC appear to show dissociation between general intelligence and moral behavior. The prefrontal lobe is where inhibition mechanisms are located, and patients with prefrontal lesions are notoriously uninhibited. All this implies that innate structures of the brain in the OFC help manage (but perhaps do not prespecify) moral behavior.

The fourth area of potential correspondence between UG and morality raises the question: Do innate pre-specified expectations cause the structure of human moralities to emerge the way they do? Can human children create moralities out of the data coming from their moral environments, or out of the raw material of pidgin-like proto-moralities? Is there a metamorality (MM), similar to the metagrammar of UG? Do developing children create a full-fledged morality out of a paucity of moral input? These are interesting and challenging questions, but evidence for MM might not find the strong support in morality creation that UG finds in language creation. First, I know of no evidence for Home Sign or pidgin type moralities being creolized into creole-moralities. Moreover, from what we observe of how cultures tend to try to teach morality to children, it seems that a paucity of moral input might result in a lack of moral development in children.

That is, whereas children seem to develop normal language despite the apparent poverty of the stimulus and their imperviousness to language correction, they might not develop normal morality with a poverty of moral stimulus, input, and correction.

However, the question seems tangible enough that we can detail an empirical scenario (which would probably still be ethically questionable). To cash out the idea, we would need child "speakers" of a pidgin morality, and we could test these children to see if they develop, on their own without guidance, a more complex Creole-morality than their caregivers. In fact, this seems highly unlikely. Rather, if a pidgin morality even exists in a community, then it is likely to be replaced by a morality, not created by the children's MM but rather supplanted by a developed religious, political, or moral system brought by proselytizers, politicians, or parents. In addition, it is interesting to note that linguistic input for a child is almost wholly positive, in that they tend to learn from information about what *is* grammatical in the language and remain relatively impervious to the correction of caregivers. That is, there is little or no direct negative evidence in language acquisition, and this has important ramifications for any meta-morality.

On the other hand, if "children typically receive only positive evidence" and "their parents do not generally correct their grammatical mistakes, then "input seriously underdetermines learning" (Ellis, 1997: 67). This means that negative evidence does not provide the necessary information for language acquisition to succeed. Moreover, in the Chomskyan view, even positive evidence cannot provide enough input for acquisition to occur, and hence the argument that abstract aspects of grammar belong to our biological endowment -- enabling us to naturally expect grammatical rules. However, in the case of moral acquisition, children not only require some negative evidence to adjust their behavior to caregiver moral standards, but they also do not always seem impervious to negative evidence and correction regarding moral behavior. That is, children can and often do respond to negative corrections from caregivers. Moreover, though not all moral instruction is negative, a lot of it still finds expression in phrases like "don't do that," or "don't do X; do Y instead." For example, eight of the Ten Commandments are couched in the

negative: "Thou shalt not..." Thus, negative evidence may perform a more effective role in moral acquisition than in language acquisition.

What does this mean for our discussion on the parallel between UG and morality? First, moral acquisition differs from language acquisition on this point, regarding the roles of positive and negative evidence. Moreover, negative evidence appears to play a more significant role in the acquisition of morality than it does in the acquisition of language. Thus, to generalize, children appear more sensitive to negative evidence regarding morality than they do regarding language. Of course, children can rebel against their caregivers, and for periods of time or even permanently, they can deny or negate the morality they have been instructed in, therefore opting for a different form of morality. (Nevertheless, they will be generally hard pressed to get along in society without expressing some kind of reciprocal altruism or Golden Rule behavior, unless they are able to survive by manipulating others). In contrast, however, it would seem virtually impossible to rebel against the acquisition of the linguistic code; learners might acquire a different dialect or style of speech than their caregivers, but they would still possess a fully recursive language. Thus, moral acquisition and language acquisition differ from each other on this point; that is, we might more naturally *expect* the rules of grammar than morality.

Perhaps MM comes from linguistic cognitive abilities related to categorization (Monroe, 2002), predicate-argument structure, (who does what to whom), or from recursion, traits that correlate well to UG, but not so much to language creation itself. In this case, therefore, the ways that predicate-argument structure and recursion enhance morality would not so much demonstrate a MM, but rather how UG and human language give us the power to invent, apply, or extend human morality itself. Later we will discuss how this power of language relates to the Golden Rule. However, a key point here indicates human beings significantly ground and elaborate morality or moral systems in linguistic ability, rather than vice versa. For example, though the OFC does not pre-specify the possible kinds of moralities that can exist, it does imply localized brain control of moral behavior. However, though UG does not directly pre-specify morality per se, it pre-specifies

the patterns of human languages, and these patterns may influence the forms of human moralities. Thus, UG may influence the form of human morality in the ways that the universal patterns of language allow for the universal patterns of morality.

5.4 Recursive Linguistic Creativity Enhances Morality

Related to language creation stands a similar concept: recursive linguistic creativity. If language creation does not directly correlate to the moral sense, what about the related idea that recursive linguistic ability enables us to create an infinite number of novel sentences? Recursion refers to the "computational mechanisms [that provide] the capacity to generate an infinite range of expressions from a finite set of elements" (Hauser et al., 2002: 1571). According to these authors, recursion may be the only characteristic that distinguishes human language from non-human communication systems. That is, other species may be able to apply recursion to other domains, such as, navigation and social relations. Nevertheless, in the human case, recursion enables humans to express their moral sense and ethical ideals about an infinite number of situations, objects, and relations. In a word, creative recursive language makes human morality creatively recursive.

For instance, we can moralize about the usage of cellular phones in public places, about humane or inhuman treatment of animals, about issues pertaining to sexuality and personhood, about the responsibility of wealthy nations to poorer nations, about dress codes, about the amount of money wasted globally each year on necktie purchases, which could be used to instead for charity, about the inappropriateness or appropriateness of different kinds of humor, about use and abuse of natural resources, about the virtue or vice of religion, agnosticism, or atheism, and we can meta-moralize about morality itself, including why we think it immoral for people to moralize about our actions.

Besides being able to moralize about an infinite number of actions, we can also moralize recursively about one act. For example, we can use the following "if/then" and "not only/but also" recursive construction. If you do not return the money you

found on the street, then the police may find out about it, and if the police find out about it, you could be charged with stealing (since this is a crime in this country), and if you are charged with the crime of stealing, then besides having a criminal record, there will be a chance that you will go to jail, and if you go to jail, then you will not be able to take care of your family, and if you cannot take care of your family, then you will not only be a criminal, but you will also be an irresponsible nincompoop of no count for putting your family into poverty, and if all these things could happen for not returning the money, then it would be better to simply return it, but if you do return it, then...

If this were not enough, we can meta-moralize not only about whether real morality exists or not, but also whether it would be ethically wrong (or right) to persuade others to follow a non-real morality. Nevertheless, the point here pertains not to whether recursion leads us to moral realism or anti-realism, but rather that (1) linguistic recursion helps us moralize about an *infinite number of actions*, and (2) it also helps us moralize *infinitely about any one act*. If this is the case, then linguistic recursion perpetually enables, extends, and enhances the range and number of animals, vegetables, minerals, people, events, situations, and even imaginary scenarios we can moralize about.

In addition to the fact that the creative and recursive nature of language makes human morality recursive, a recursive moral code also stands as a uniquely distinguishing feature of human morality compared to the proto-morality or altruism of non-human species. That is, degrees of recursive ability between species will differentiate the degrees of moral ability between species. Hauser, Chomsky, and Fitch (2002) say examples of animal recursive ability (navigation, number, and social calculus) stand as potential precursors to recursive language, and they suggest that domain-specific aspects of recursion became domain-general in humans. Along these lines, humans can combine recursive abilities; thus, recursive social calculus relates to how recursive language helps humans possess a recursive theory of mind. Thus, with language one can think: "I think that Henry thinks that Kenny borrowed Jim's book and should return it lest he fall out of favor with Jim and the rest of us." This stands as a linguistic form of social calculus that

demonstrates moral differences between species. For example, how might apes express a recursive theory of mind? If Chimp A thinks that Chimp B and Chimp C are in conflict with each other, and if Chimp A attempts to help B and C reconcile, this behavior might stem from a recursive theory of mind. However, the important point here concerns how recursive language extends other recursive abilities in the moral realm -- for example in how we think about and attempt to reconcile numerous people involved in a conflict.

Regarding chimpanzee reconciliation (de Waal, 1982; Arnold and Whiten, 2001), recursion, and theory of mind, we cannot easily substantiate the claim that apes can read mental states (Povinelli and Vonk, 2004; Premack, 2004), which they could employ when reconciling. Moreover, though there may be some cases where chimpanzees can know what conspecifics know and do not know (Hare et al., 2001), whether they have the ability to attribute states of mind remains a controversial, complex, and debated point (Arnold and Whiten, 2001). Basically, this is because, "there is no easy way of making an a priori transition from behavioral similarity to psychological similarity" (Povinelli et al., 2000: 27). Interestingly, Povinelli, Bering, and Giambrone propose...

... that the majority of the most tantalizing social behaviors shared by humans and other primates (deception, grudging, reconciliation) evolved and were in full operation long before humans invented the means for representing the causes of these behaviors in terms of second-order intentional states. (Povinelli et al., 2000: 25)

If this is true, then higher order representational abilities such as recursive language would add a whole new array of behavioral repertoire to the organism on top of these already existing behaviors.

More importantly for this discussion, language stands as a primary means to access the mental states of others, for though I may be able to deceive another about my intentions, I can also make my real intentions known. Moreover, I can tell you what I think you are thinking, and you can tell me whether I am more or less correct or not, or I can tell you what I think you think Frank is thinking, and you can tell me whether you think I am more or less right or not. Therefore, if language does not make possible a recursive theory of mind, at least language greatly

extends it. Thus, no matter what ultimately causes apes to reconcile, linguistic recursion and recursive theory of mind will greatly extend this behavior in human beings.

For example, you may be a noisy neighbor, and you may not know that your noise bothers your neighbor, but your bothered neighbor could solve this problem directly by talking to you, or she could recursively communicate with you through another neighbor. She may tell another neighbor of the problem, and ask him to approach you with a request to be quieter. When she does this, you can apologize to her through the mediator without even seeing or speaking to the offended party, and recursive language will expand on this recursively social behavior. For example, US President George Bush made a recursive social apology for Iraqi prisoner abuse through King Abdullah of Jordan while using recursive language.

Describing his conversation with the King, Mr Bush said: "I told him I was sorry for the humiliation suffered by the Iraqi prisoners and the humiliation suffered by their families.... I told him I was equally sorry that people who have been seeing those pictures didn't understand the true nature and heart of America" (The Scotsman, 2004: May 7).

The above exemplifies how humans use recursive language with a theory of mind to extend the range and variety of human moral behavior. Regarding recursion, Aitchison (1999: 79) says, "we can never make a complete list of all the possible sentences in any language," and this suggests that we can never make a complete list of all the things we can moralize about. In short, recursion stands as a defining feature of human language and social calculus with its linguistic access to other minds, which strongly affects human sociality and morality.

5.5 Creativity, Naming, and Morality

In addition, a building block of recursion, "the naming insight" also extends and expands the human ability to moralize. Speaking of the origin of human language, Aitchison (1999: 19) asserts that besides being able to produce a range of sounds, humans "must have attained the 'naming insight,' the realization that sound sequences can be symbols which 'stand for' people and objects." Non-human

species such as some primates have the cognitive abilities to name things (Savage-Rumbaugh et al., 2001), and other animals, such as dogs, have the ability to recognize names for things (Kaminski et al., 2004).

Whales and dolphins have signature calls that indicate their presence to the group, and vervet monkeys have a number of apparently innate calls regarding predators. However, we generally see extensive name-production or name-recognition in non-human species because we either teach "names" to these animals or because these animals have had intensive interaction with humans. Moreover, for animals that do possess minimal naming insight, the question arises as to whether they use it to attribute moral values to named items. Do language-trained animals name objects with a moral sense of good, bad, right, and wrong? This is an extremely interesting question. For example, the chimpanzee Lucy was taught American Sign Language, and researchers purportedly observed her lying, apologizing, and thus experiencing guilt. In one conversation, the researcher Roger Fouts (1998: 151-156) questions Lucy about excrement on the floor.

Fouts: WHAT THAT?
Lucy: WHAT THAT?
Fouts: YOU KNOW. WHAT THAT?
Lucy: DIRTY DIRTY.
Fouts: WHOSE DIRTY DIRTY?
Lucy: SUE.
Fouts: IT NOT SUE. WHOSE THAT?
Lucy: ROGER!
Fouts: NO! NOT MINE. WHOSE?
Lucy: LUCY DIRTY DIRTY. SORRY LUCY.

We cannot extrapolate too much from this conversation, but it is suggestive of the role language might play in affecting the moral-like behavior in naming enabled non-human species who have been influenced by humans. Nevertheless, we can still say that the way humans use the naming insight stands as a distinguishing feature of human morality. Not only can we name objects, people, events, and concepts, but we can also coin new names for anything, and most importantly, we can attribute the values of good, bad, right, and wrong to virtually all the things we name.

Hence, because we can name stuff, we can moralize about the things we name in a very simple and protolinguistic fashion. For example, "monogamy is good; "polygamy and polyandry are bad;" or for those opposing the legalization of marijuana: "weed is bad;" or for those in favor of trickle down economics "greed is good." Such moralizings are relatively simple because they do not require recursion, syntax, and argument structure; that is, we can say "weed bad" or possibly "bad weed" to get our message across. Moreover, argument structure "who does what to whom" does not come into play in these phrases. Thus, we can moralize protolinguistically, with simple labels and without argument structure.

Can non-human species -- who recognize or produce names -- moralize about the things they name? I suggest that they cannot moralize on their own about even the most basic things. However, under human cognitive and linguistic influence, these animals probably can attain some sense of right and wrong about the things they name-recognize and name-produce, which, for example, could obtain through operant conditioning. However, it is likely that this will only happen because humans first attributed good and bad to particular names, through punishments, rewards, and verbal affirmations and reprimands. Nonetheless, if we inquire into animal ability to ascribe values of good and bad to objects, events, or behaviors, this might produce interesting and testable questions about the moral capacity of name-enabled species.

With all this talk about how naming ability and language enhance morality, the skeptical reader might wonder how we might use language for immoral purposes. Thus, before moving on, a small caveat is needed. For example, a large literature exists on the human ability to deceive with language (Renshaw, 1993; Stiff and Miller, 1993; Wortham and Locher, 1999; Galasinski, 2000; Meltzer, 2003; Newman et al., 2003). Hence, though language has the power to extend moral behavior, it also holds the opposite power to deceive others, negate morality, and advance malevolence. Thus, language may give us the ability to create an alternative morality, such as in George Orwell's novel *1984*, in which "Newspeak" is used to teach, "War is peace. Freedom is slavery. Ignorance is strength" (Orwell, 1950: 7). The topic of how language can facilitate anti-altruism and immorality

transcends the focus of this thesis. However, though we must acknowledge the negative power of language to deceive and serve selfishness, as well as the cost of deception and the punishment of deceivers (Hauser, 1992), this does not negate the positive power of language to enable, extend, and maintain human altruism and morality.

5.6 Linguistic Displacement Enhances Morality

In addition to how naming ability helps us assign moral values to what we name, language also helps us make abstractions, and this highlights the unique feature of human language called displacement. Crystal (1992: 26) defines displacement as the ability "whereby language can be used to refer to contexts removed from the immediate situation of the speaker (as in the cases of tenses which refer to past or future time)." Animal calls, on the other hand, only refer to "specific situations, such as danger and hunger, and have nothing comparable to displaced speech" (Crystal, 1992: 26). Hence, displacement enables humans to refer to things removed in space, time, and even reality from the speaker, referencing the hypothetical or unreal. Though some species exhibit limited displacement ability, as in bee dancing, this still refers to the specific physical location of displaced nectar. Thus, displacement exhibits unique features in human language that transcend concrete situations.

Hence, how could linguistic displacement uniquely enhance and extend human morality? For one thing, as previously mentioned, it enables us to moralize about the past and the future, and though some animals might feel regret about past events, such as an elephant mourning the loss of kin, this is still quite different from moralizing about past events. Is it possible that two bonobo chimps could be made to regret their secretive copulation through a verbal rebuke even if the dominant male who might physically oppose such behavior never found out about it? Would it be possible through verbal or any other means to make a male elephant mourn the death of conspecifics he has not actually physically seen? However, even a human child in the first grade of elementary school can reflect on a parent's

scolding: "it was not good that you lied to your teacher, telling her your dog ate your homework, instead of the truth that you simply forgot to do it."

Besides past-event-moralizing, with language we can turn our attention to the future and instruct a child in the following way. "Tomorrow you will apologize to your teacher, and tonight (future-displacement) you will write your ancestors (abstract-displacement) an apology, reflecting on how you can remember your homework and reasons why (hypothetical-displacement) you should not lie again (future-displacement)." Besides moralizing about the past and future, displacement enables us to moralize about the hypothetical and unreal. For example, "if your boss pressured you to lie about your company's financial accounting, would you go along with your boss or blow the whistle on him? Moreover, in a practical ethics course, participants can discuss ways to carefully deal with thorny ethical issues before they ever encounter them. Additionally, we can think about fictional or futuristic ethical dilemmas. If you suddenly found yourself with the ability to foresee the future with 80% accuracy, and the government asked you to predict terrorist activity and arrest "pre-crime" terrorists before they can act, what would you do about it? In short these examples show that displacement, as a defining feature of human language also distinguishes human moralizing from proto-moralities because it enables us to think morally about that which is removed from us in space, time, and even reality.

Along these lines, it is interesting to note the relationship between displacement and recursion. First, displacement does not require recursive embedding, for we can refer to the future, the past, places, and non-realities in proto-linguistic ways (with 1-word utterances): tomorrow, yesterday, Venus, Mars, Hercules, and Zeus. Incidentally, though we can name these concepts in 1-word utterances, we may need recursive ability to understand at least some of them. For example, even if we see statues or images of the god Zeus (upholder of justice and morals), we still cannot understand what the name means without a recursive explanation. Nevertheless, though displacement does not require recursion, with recursion, displacement becomes unlimited -- enabling us to moralize without end about anything removed from us in space, time, and reality.

5.7 The Influence of Stimulus Freedom on Morality

Not only can humans use language liberated from space and time, but we also can create language output at any time in the absence of any external stimuli. That is, we are not only free to linguistically displace, we are also not bound to events or people in order to talk about them. This characteristic of "stimulus freedom" (Chomsky, 1959) is the flip side of displacement. (Displacement is a feature of the language, and stimulus freedom a feature of the speaker). How does this characteristic of human language relate to morality? For one thing, moral stimuli, people, events, situations, do not need to be present for us to moralize about them. Hence, stimulus freedom, as a cause, characteristic, or function of displacement enables us to talk at any time about any past, present, future, real, or unreal stimuli that we charge with moral meaning. Therefore, with stimulus freedom, we can reflect on any past moral event, or we can anticipate any future moral event. This kind of reflection underscores and reiterates the fact that when we talk about past morally charged events, we can prepare and respond to them before we experience them.

Moreover, this supports the idea that in this freedom from moral-oriented stimuli, we can also foresee heretofore not-experienced moral events, and contemplate how we might respond to them if they were to come upon us. Such behaviors, of course, would happen as a function of linguistic displacement, but looking at them from the perspective of stimulus freedom also adds another subtle nuance to their expression. The difference is this: displacement refers to a *characteristic of the language*, which for the purposes of our discussion express moral concern regarding places, times, events, and worlds removed from the communicator. Stimulus freedom refers to a *characteristic of the language user*, who when discussing morality can *choose* to moralize at any time in the absence of the morally charged stimulus. Setting aside philosophical issues concerning freedom and choice, stimulus freedom is a psychological function that enables us to communicate independently of stimuli, and when we combine this with linguistic

recursion, we displace with language. In a word, we can linguistically displace our moral thinking while free from morally charged stimuli.

Related to stimulus freedom stands Bickerton's (1996) distinction between online and offline thinking. Online thinking forms automated responses to sensory information, including objective knowledge of our environment and subjective knowledge about our inner states of mind. Online thinking also forms a kind of stimulus-response behavior (as opposed to stimulus free) that we use, for example, when we flee from a predator. In contrast, offline thinking coheres with stimulus freedom and refers to, for example, how Ichiro Suzuki, the famous Japanese baseball player, reviews every play, every success, and every failure after every game while he polishes his shoes and cleans his glove. Bickerton claims syntactic language makes possible this kind of offline cognition. Thus, besides behavioral review, language enables us to practice for future events, like the social worker that practices how she will advise at-risk clients who suffer from drug addiction or who practice sexually risky behavior. This practice can range from full "dress-rehearsal" to pure mental rehearsal. Thus, language enables us to rehearse moral offline thinking -- to physically role-play or mentally visualize altruistic acts.

For this discussion, it is important to note that linguistically enabled offline thinking appears to be unique to humans. Animals, for instance, do not *directly* practice their *actual* killing or fleeing techniques in the absence of prey or predators, though play does appear to serve a similar role. However, we humans will not only intensely think through and practice any number of behaviors, but we will also employ various means for offline cognition about morality. For example, doctors attend ethics classes or seminars; schools in many nations have moral education programs; and millions of people voluntarily receive religious instruction that at least partly serves as offline thinking about morality. The point: language appears to strongly enable or determine offline thinking that is stimulus free. Thus, humans can universally employ linguistically enabled offline thinking in order to practice and refine their moral and altruistic responses to anticipated situations.

The reasons why humans possess stimulus freedom and linguistic displacement point to interesting questions that merit further research. However, these

characteristics clearly distinguish us morally from other species and their substrate moralities. That is, a great deal of moral interaction takes place in the presence of moral stimulus while using language about the here and now, "No Johnny! Don't poke Mary in the eyeball!" On the other hand, moral systems clearly take advantage of stimulus freedom and displacement in a myriad of ways. Here are some examples: Aesop's Fables give practical moral lessons nevertheless based in unreal and impossible events. Many religions teach their followers that an invisible God or gods carefully watch and evaluate their behaviors. Moreover, these divine beings temporally or eternally reward and punish people for their good and bad behaviors. Thus, these simple and common examples demonstrate that humans actually practice offline moral reasoning -- enabled by language -- in the real world.

In addition, stimulus free language users have created abstract laws and constitutions, and these texts express morality in displaced language, delineating what is right, lawful, and good for whole nations. In short, the moral world as we know it cannot exist without displacement and stimulus freedom. Morality requires stimulus freedom, and our ethics entail linguistic displacement. In this way, as the design features of human language radically extend human morality, the question emerges whether displacement and stimulus free moralizing can actually help produce moral or altruistic people. This thesis may only partly answer this question; however, humans clearly employ stimulus freedom and displacement when moralizing and attempting to apply their moral ideas to concrete situations. Moreover, we can safely assume that human beings experience varying degrees of altruistic success (and egoistic failure) when we utilize these design features of language in the moral realm.

5.8 Cultural Transmission and Moral Values

In addition to the design features mentioned above, cultural transmission stands as another design feature of language that extends human moral ability. Smith states: "With the exception of birdsong, all other non-human communication systems are taken to be genetically transmitted" (2003: 4). However, though genes transmit the

seeds of grammar (UG) between generations, human language also transmits the flora of culture between generations. Moreover, as language transmits culture, it also transmits moral culture. Steven Pinker states:

A common language connects the members of a community into an information-sharing network with formidable collective powers. Anyone can benefit from the strokes of genius, lucky accidents, and trial and error wisdom accumulated by anyone else, present or past. (Pinker, 1994: 16)

This information-sharing network includes the data of a collective moral culture. Hence, for example, moral traditions inherent in democracy, communism, Confucianism, Christianity, or Islam exist as moral cultures that only humans with language can create and transmit between generations. In short, the linguistic transmission of moral culture adds another significant foundation stone to the house in which humans live as uniquely moral creatures.

5.9 Language, Categorization, and Morality

Concerning how features of language enable human morality, language also supports our ability to categorize, and this stands as a vital aspect of morality. This was mentioned in the discussion of moral realism; unfortunately, according to Harnad, "Very little is known yet about the brain mechanisms of category perception and learning" (2003: 3). Therefore, our incomplete understanding about categorization will limit what we can say about language, categorization, and morality. Nevertheless, besides the innate aspects of categorization rooted in sensorimotor experience, such as face recognition, color differences, and categorical perceptions of speech sounds, we want to know how language might help to influence, inform, or even create our perceptions of moral categories.

According to Lane (1965), Lawrence (1950) and replicated in Goldstone (1994) cited in Harnad, (2003), learning on its own can induce categorical perception. Moreover, as Harnad points out, thousands upon thousands of lexical categories logged in dictionaries could not possibly be innately grounded in consciousness, in spite of the extreme claims of some nativist theories. Therefore, these categories must be learned, and this raises interesting questions regarding the role of language

in category creation and categorical perception (CP), which Harnad states in this way:

Some of our categories must originate from another source than direct sensorimotor experience, and here we return to language and the Whorf Hypothesis: Can categories, and their accompanying CP, be acquired through language alone? (Harnad, 2003: ¶ 21)

Harnad mentions neural net simulations (Cangelosi, 2001, cited in Harnad, 2003) that suggest when the mind grounds a set of category names in sensorimotor experience, it can translate them into higher order categories through Boolean combinations, e.g., (man = male & human) and (bachelor = unmarried & man). However, Harnad concludes that we have yet to experimentally demonstrate "language-induced CP-effects" in human subjects. Nevertheless, he suggests that the evidence for sensorimotor and learned CP (as opposed to language-induced CP) still demonstrates the "Whorfian power of naming and categorization, in warping our perception of the world" (2003: ¶ 23). In this he contrasts a "full-blown language effect" as opposed to a "vocabulary effect," and therefore if language can warp our perceptions of the world, we need to demonstrate two things: (1) how the way we name things affects perception and (2) how what we are instructed about these names affects perception.

This is interesting because it raises a number of difficult questions. What is the difference between a vocabulary effect and a full-blown language effect? What are the differences between learning-induced CP-effects and language-induced CP-effects? How might learning-induced CP-effects and language-induced CP-effects be related? Does the way we name things and what people teach us about certain name-categories create a language-induced CP-effect? It only seems obvious that what I tell you about X will influence your perception of X and the world that X belongs to because X must relate to the world in numerous ways. If this is the case, then the problem remains that we simply need to demonstrate empirically what seems apparently obvious.

Nevertheless, despite reservations and questions, Harnad (1987) concedes that language may uniquely mediate CP because it gives us the labels and the

descriptions that subtend most of our categories. Thus, for Harnad, language enables us to name and identify positive and negative examples of a category from an array from similar or easily confused options. Thus, when we categorize correctly particular instances based on invariant features inherent in the given cases, Harnad claims that we associate the category name with the sensorimotor categorical representation. The category names provide the linguistic symbols that enable us to learn by verbal description. Thus, language adds a powerful descriptive force to sensorimotor perceptions, reinforcing and adding to our categorical perceptions of empirical experience; moreover, language enables us to learn about displaced and abstract entities and events purely by description.

The point here concerns how language may subtend moral categories. To paraphrase Harnad, language might mediate moral categorical perception because it gives the names and descriptions that underlie our moral categories. The idea of moral categorization also relates to the putative semantic primitives such as "good" and "bad" as suggested by Wierzbicka (1996). We do not want to overstate our claims about semantic primitives; nevertheless, Wierzbicka soundly asserts that the concepts of "good" and "bad" are conceptual categories that are "innate and fundamental elements of human thought" (1996: 52).

Wierzbicka claims that all languages have words for "good" and "bad," and the universality of these concepts provides evidence that they stand as semantic primitives. She even states that experience can teach us to regard certain things as "good" or "bad," but it cannot teach us the concepts of "good" and "bad." This is important for our discussion here because if "good" and "bad" exist as fundamental, innate, and unlearned categories, then language may enable us to put things in those categories. This should be so, even if "good" and "bad" are not semantic primitives. Thus, we can accurately or inaccurately, categorize behaviors or people as good or bad, and this will have serious implications for morality. The following anecdote highlights the moral implications of categorization -- especially miscategorization.

Imagine a country field in the rural south of the US. Hear the sounds of children laughing. It is late summer, early September, and about five or six elementary first

grade children of various racial backgrounds are playing together after one of their first days at school. The Wallace family drives up in their old white Ford Fairlane to pick up their child. The child, Wally, hollers "Goodbye" to his playmates and gets into the back seat of the car. As the Wallaces drive off, Mrs. Wallace turns around and asks: "So Wally who were you playing with?" Wally answers, "My new friends from school." Then Mrs. Wallace says: "Well, you know we'd rather that you play with *your own kind*." Wally replies: "What do you mean Mommy? We're in the same class." But Mrs. Wallace shakes her head. Then Mr. Wallace cuts in. "Them weren't kids you was playing with. They're more like animals than humans, and we don't want you playing with un-humans. They are bad.

Wally cringes at first, but out of respect and fear of his father, he says: "Yes, sir." Over the years, Wally's parents continue to verbally reinforce to Wally their categorical perception of race, about who deserves fully human treatment and who does not. In time, Wally begins to think the same way, and when he grows up, he teaches his children the same things. He evaluates the intentions and actions of those he perceives as less than human always with prejudice and distrust, thus tainting his every interpretation of their behavior, and he tends to only interact with those people who share his prejudices. This example reiterates points made already; however, the additional argument here is that as a linguistic species we may possess fundamental and innate semantic categories such as good and bad. Thus, as we use language to help us interpret and guide our experience, we will accurately or inaccurately place events, behaviors, and even people in these naturally occurring categories.

The existence of a minimal set of semantic primitives has significant implications for this study. Though we should be cautious, some scholars claim "the prospects for a universal semantic metalanguage remain bright" (Goddard, 1997: 207). We can see this in a growing body of research on the topic, for example, (Wierzbicka, 1992; Goddard and Wierzbicka, 1994; Goddard and Wierzbicka, 1994; Wierzbicka, 1996; Goddard, 1997; Onishi, 1997; Goddard et al., 2002). Additionally, without the existence of some semantic primitives or near primitives, we would have no "constant and language independent points of reference"

(Goddard and Wierzbicka, 1994: 81) between languages. Thus, if the semantic primitives of good and bad are real (as they appear to be), this supports the idea that we cannot easily avoid categorizing things, behaviors, or events as good or bad. We cannot easily circumvent moral categorization, and even if we call something "neutral," this actually slides over to the category of "good" because it is obviously "not bad." It is as if the innate primitives of good and bad wait for us to employ language as the category filter, which we use to place "things" in these extant categories.

Hence, the above is an all-too-real scenario of how humans can employ language to warp the moral categories we initially perceive in a sensorimotor or empirical fashion. Wally would see his friends as just like him, people with two eyes, two hands, and two feet, and who feel, think, and act in ways similar to him. For all he can see, they belong in the same category that he is in, which he thinks of as good. For Wally, he could learn racism in three ways: (1) observe and imitate the behavior of racists; (2) feel it physically and emotionally if he is the brunt of racist abuse; and (3) learn racism through verbal interaction and categorization. We may find it difficult to ascertain the causes of racism. However, we will also find it hard to imagine racism *not* happening on grand scales (as in American Slavery, South African Apartheid, and Nazi Germany) without people using verbal cues to warp the perceptions of the masses to maintain and perpetuate racism.

That is, to make such a system work, people in power will use language to call a "spade" a "hammer," or they will use another linguistically embedded argument to show why a particular group deserves varying degrees of "inhumane" treatment. As already mentioned, Nazi Germany used the term "untermenschen" or subhuman to redefine a whole race of people. Incidentally, if racism can be taught through language, then we can reinforce human equality through language also. However, the key point here concerns the idea that the concepts of good and bad appear to be real semantic primitives or near primitives. Nonetheless, due to the influence of language, we are able to accurately or inaccurately categorize our experience and so place that which is actually bad into an innately existing category of good, or take that which is good and somehow place it in the semantically primitive

category of bad. Hence, our ability to lie or tell the truth to ourselves (using language) regarding the apparently real categories of good and bad has overwhelming implications for human morality.

5.10 Design Features of Language, Identity, and Morality

This brings us back to Monroe's (2002) discussion on the "importance of identity and categorization for moral action." Monroe thinks that UG provides an innate and universal basis for categorization.

Children are born with an inherent capacity to rationally organize sensory information into discrete and manipulatable symbols (words) which can be recombined according to a finite set of rules (grammar) into a potentially infinite number of sequences. (Monroe, 2002: 503)

This is important for Monroe because although she argues that identity forms a basis for altruistic behavior (more than rules do), UG provides an indirect basis for universal morality because it helps us translate moral principles across cultures.

That is, if UG helps delineate universal categories, such as argument structure, then the basic moral principle of the Golden Rule, for example, may also find universal translatability and comprehensibility. Incidentally, the phrase "Do to others like you want others to do to you" consists *only* of items from the inventory of semantic primitives (Wierzbicka, 1996) that may likely be lexical universals "in the sense of having a counterpart, an exact translation, in every human language" (Goddard, 1997: 1999). Nevertheless, this kind of potential universal morality still differs from a universal grammar, and the similarities and differences between UG and morality that we have already discussed. At this point it seems difficult to make a claim for a universal morality that completely parallels the innate principles of UG because generativists have not yet empirically found them. However, if a universal morality does exist, it could probably only emerge through categorical, syntactical, and semantic abilities afforded to us by language, and it would probably be expressed in semantic primitives.

Nevertheless, the worst moral disasters of human history all seem to share the same common denominator: the oppressed "have all been categorized as less than human" (2002: 504). Hence, Monroe emphasizes that, "people's categorization influences their treatment of others." That is, we use language to categorize how we treat people, and it is difficult or harder for us to mistreat those people we place in the categories we see ourselves in. To put it another way, it is perhaps easier for us to treat people well when we categorize them as equal to us, and we cannot escape using language in order to accomplish this. Hence, even though Monroe does not strongly support the view that we can develop altruistic persons by inculcating them with rules and laws (based in language), she still considers language as absolutely essential for growing altruism as we form our identities and categorize others. Thus, this theory makes the empirical claim that how we morally treat others will highly correlate with how we verbally categorize them, as "others who are people just like us." Conversely, immoral treatment of others will correlate highly with the verbal classification that the "others are subhuman." Incidentally, Monroe emphasizes identity and categorization in moral development -- and that identity supercedes choice. However, when we use language to teach morality -- through rules, laws, and or spiritual values -- this may still have significant influence on moral development. That is, linguistically based morals may still affect how we develop our sense of identity and categorization and play a constant and important role in how we go about making moral choices.

Monroe's approach raises interesting questions for the focus of this chapter regarding the properties or design features of language that make human morality possible. For example, how does syntax enable categorization? What role does argument structure play in helping us develop a sense of self and affording that selfhood to others? Pinker and Bloom (1990) refer to the language of thought or mentalese and claim that the propositions in this representational system are symbols that represent people, objects, events, and their categories. Thus, at the root of language, we find the cognitive structures that enable us to categorize entities, actions, events, and behaviors. If so, this also means that language enables us to classify other people as human or non-human, as equal or unequal, or simply as good or bad. If this is true, humans require language or the cognitive abilities

that subtend language in order to develop an altruistic personality. *That is, morality fully requires language.* Language enables us to form our identity and afford the kind of identity to others that merits altruistic (or egoistic) and moral (or immoral) treatment. Moreover, language not only helps us develop identities and appropriate moral categories, it also allows us to attribute various meanings and values to the categories we create.

The above discussion raises two important questions. First, what roles do the design features of language play in human identity development? Second, how do these linguistic features enable virtual relatedness? For the first question, as mentioned previously, design features of language contribute to human identity in many ways. For example, linguistic differences help distinguish the identity between ethnic groups, and recursive linguistic ability distinguishes humans from non-human species (Hauser, et al., 2002). Moreover, the advanced semanticity of human languages -- the ability to match specific signals to specific meanings also facilitates human identity development. Hence, the identities we linguistically attach to fellow humans and other conspecifics have moral implications because the meaning we attach to them affects how we perceive their identity and thus how we behave towards them and value them.

5.11 Design Features of Language and Virtual Relatedness

Now to conclude this chapter, we will look at how the design features of language enable virtual relatedness, *vr*. Virtual relatedness suggests that linguistic expressions of the Golden Rule "do as you would be done to" can influence us to treat non-kin more as if they were our kin. It is interesting to look at the Golden Rule from the perspective of natural selection. Selection pressures us to look out for our own reproductive fitness; however, the Golden Rule implicitly embraces the idea of selection operating on behavior: "you naturally look out for your own genes, but it is also good to treat the genes of others as if they were your own. Thus, the Golden Rule encourages us to act in ways over and above the "call of selective duty." We have previously discussed numerous issues raised by virtual relatedness; however, important discussion remains concerning how the design

features of language support *vr* as a central and universal aspect of morality and altruism.

Regarding UG, which prewires us to expect and find the basic principles intrinsic to all languages, such as argument structure, it seems that UG helps us understand and therefore experience the Golden Rule. This is different from being innately predisposed to a particular universal ethic. First, if we assume the Golden Rule is good, we still have the ability to redefine who benefits from it; thus, we can categorize people outside of the reach of the Golden Rule by calling them subhuman and unworthy of it. Second, this implies that how we categorize and identify people (using language) determines how far we extend the Golden Rule. Third, therefore, we cannot say that UG predisposes us to a universal reciprocity ethic or Golden Rule.

However, that UG naturally allows us to experience the Golden Rule simply means that UG enables us to think in terms of virtual relatedness. Though this is a reiteration of previous points, we need to clarify the rather subtle relationship between UG and *vr*. If we try to express the Golden Rule in protolanguage, we begin to see that *vr* would probably not exist except in a language formed by UG. "Do good to others as you would have them do good to you." This phrase relies only on semantic primitives (Wierzbicka, 1996) (See Figure 5.1), argument structure (who does what to whom) and recursion (two phrases embedded into each other). Thus, we cannot express this in two-word, non-syntactical, non-combinatorial, protolanguage. In a word, because of the relations implicit to the Golden Rule, we may not be able to express it without the kind of innate categories and expectations that UG provides.

In addition, though the Golden Rule and *vr* require syntactic recursion, we also can express the GR and *vr* more widely through the design feature of linguistic creativity. We can say: "do unto fellow countrymen as you would be done to;" "do unto fellow Europeans as you would be done to;" "do unto fellow humans as you would be done to;" and "do unto all living things as you would be done to." We can therefore expand or collapse the boundaries of virtual relatedness as far as we want through creativity. Moreover, the concept also appears to depend on the feature of

displacement. Do unto other people (who are not necessarily present now) as you would be done to (in any hypothetical situation that is also removed from you in time, space, or even reality). Hence, we can employ the concept of virtual relatedness in the absence of any moral agent or event, and so the Golden Rule also relies on stimulus freedom -- as an aspect of the language user.

Figure 5.1: The Golden Rule and Semantic Primitives

<i>Do good to others, as you would have them do good to you.</i>	
--	Words in the GR that belong to the basic set of semantic primitives
Actions	Do
Attributes	Good (not bad)
Determiners	Others
Similarity	Like, as
Substantives	You, someone, person, people (assumed in "others").
Note:	The rest of the phrase relies on argument structure, recursion, and hypothetical displacement.

In addition, since the Golden Rule finds expression in virtually all the world's religions and cultures in their proverbs and sacred books, and it is also possible to teach this concept to conspecifics, it is therefore also clear that the Golden Rule takes advantage of another important feature of language: cultural transmission. Virtual relatedness calls people to transcend their biological imperatives inherent in genetic selfishness and behave as if they were related to non-kin. Therefore, if we can express virtual relatedness as a real behavior, then this requires a mechanism that can outpace the pressures of natural selection. Cultural evolution is a process that outpaces natural selection, and since we teach *vr* through cultural transmission, this helps facilitate behaviors that outpace genetic transmission -- behaviors that we describe by the concept of virtual relatedness.

Lastly, it is interesting to note that the Golden Rule concerns both identity and categorization. Basic to this idea is our own identity and how we want to be treated; we then put another person in the same category as ourselves, and we offer them the treatment that we also desire for ourselves. Thus, the concept of virtual relatedness, a universal element of human morality, cannot find expression, cannot even be understood, without the design features of language outlined in this chapter. Moreover, the altruistic person who expresses *vr* (Monroe, 2002) must do

so through a certain kind of categorization and identification -- cognitive and behavioral processes that require language. In sum, the design features of language appear basic to human moral impulses, as seen in the Golden Rule; moreover, the fullest forms of human altruism and morality expand the Golden Rule to wider and wider circles of conspecifics (Singer, 1997). This expanding process requires us to recategorize and reidentify conspecifics within the circle of our altruistic behavior, and as we have shown in this chapter -- language greatly enables this process.

CHAPTER 6

THE INFLUENCE OF LANGUAGE ON MORAL PERCEPTION

6.1 Linguistic Relativity Light in the Realm of Values and Morals

As previous chapters dealt with how categorization, identity, and language design features help support and extend altruism and morality, this chapter discusses the influence that language may have on moral perception and categorization. Here we look at how language influences thought (the extent to which language and thought are mutually interdependent), a controversial idea, assumed to be discredited, but one that is experiencing a renaissance in a lighter, more nuanced, and rigorous revision (Lucy, 1992; Gumperz and Levinson, 1996; Boroditsky, 2001). Though linguistic relativity may seem incidental to the topic of language and morality, nevertheless these themes subtly relate to each other. Up to this point, we have seen that language can help facilitate altruism and morality in humans to the degree that altruistic and moral behaviors could actually extend beyond the limits of selfish genes. If language can influence moral behavior and thinking in this way, then we may also ask how linguistic differences might also subtly affect differences in moral thinking across cultures. That is, if language in general can influence moral and altruistic behavior and thinking, can differences between languages differentially affect altruistic behavior or thinking in speakers of those languages? In short, if language influences morality, then linguistic variation between languages may also differentially influence morality across cultures.

Evidence for cross-linguistic influence on moral perception and categorization would be interesting in its own right, and it would lend support to the thesis that language plays an essential role in the origin, evolution, and continuance of human morality. This is because demonstrating linguistic influence on moral perception would show an additional power that language wields over morality. Many linguists are rightly skeptical of linguistic determinism, that language determines the way a group of speakers think, or even the softer idea of linguistic relativity, that variation among languages shapes variation in thought. To quote Steven Pinker (1994: 57-58): "It is wrong, all wrong," and here is the reason why: "there is no scientific evidence that languages dramatically shape their speakers' ways of thinking." Nevertheless, due to improved hypotheses and new (but still somewhat controversial) evidence, some researchers are rethinking linguistic relativity (Lucy, 1992; Gumperz and Levinson, 1996; Niemeier and Dirven, 2000; Pütz and Verspoor, 2000; Boroditsky, 2001). This research compares languages at key points of difference and looks for subtle as opposed to dramatic influences on thought. We can call the approach "linguistic relativity light."

According to Lucy, empirical research in linguistic relativity needs to incorporate the following theoretical and methodological principles:

The approach advocated here emphasizes that research on the hypothesis must be based on a comparative analysis of two or more languages, that differences between the languages should be used to generate hypotheses about possible and non linguistic behavioral consequences, and that these hypotheses should be tested by careful comparative assessment of the individual behavior of speakers of the various languages. Further, it has been argued that such research will proceed most effectively at present if both the linguistic comparison and the language-cognition linkage are anchored in terms of linguistically defined referential categories. (Lucy, 1992: 149)

Along these lines, the experiment described in this chapter (1) compares two languages: Japanese and English (J/E); (2) suggests how linguistic differences may influence non-linguistic behaviors related to morality; and (3) tests these hypotheses regarding the behavior of J/E bilinguals with subjects responding to the same moral dilemmas in both languages to see if we can isolate differences in responses to differences in language. These three points frame the study discussed below.

The goal of this research should test whether linguistic differences will influence speakers of each language to perceive, remember, attend to, or even respond differently to moral dilemmas as they are distinctly encoded in each language. This corresponds with Crystal who says:

A weaker version of the Sapir-Whorf hypothesis is generally accepted. Language may not determine the way we think, but it does influence the way we perceive and remember, and it affects the ease with which we perform mental tasks. Several experiments have shown that people recall things more easily if the things correspond to readily available words or phrases. And people certainly find it easier to make a conceptual distinction if it neatly corresponds to words available in their language. (Crystal, 1997: 15)

Crystal's balanced, open-minded, and cautious point of view is helpful because the ideas of "conceptual distinction" and "available words" stand as potential areas for research regarding how language might shape moral thought.

6.1.1 Justifying Research into Linguistic Relativity and Moral Perception

Regarding linguistic relativity and moral perception, we can suggest two ways to justify this controversial research. It is controversial because of its perceived lack of empirical validity, as Pinker mentions above; moreover, academics have shunned it because linguistic relativity might suggest that one language is better than another. For example, strong determinism might suggest that one language causes its speakers to not see a particular color, or regarding morality someone might say that Language A gives its speakers a better morality than the speakers of Language B. Generativists may oppose linguistic relativity because it appears to contradict the idea that language is a universal human trait and a great equalizer. This study avoids these extremes by looking for subtle and indirect influences that linguistic differences might have on moral perception. Moreover, this chapter presents no prescriptive claim that one language or moral system is better than another, but only explores the idea that linguistic differences might differentially affect moral development, perception, or attention across languages and cultures.

We can further justify this study by recalling that language serves as a mechanism for moral development; therefore, cultural value and moral differences may also partly emerge through linguistically guided ontogenetic processes. However, this is

not a claim that distinct grammars or lexicons create value and cultural differences but rather that linguistic differences might shape moral development, attention, and perception in subtly different ways. With this in mind, we can see numerous reasons why it seems appropriate to compare English and Japanese. First, except for recent borrowing, they are historically unrelated languages -- very different from each other. Second, Japanese culture and western English speaking cultures also tend to express distinct values. Third, many J/E bilinguals intuitively speak of the differential influence each language exerts on their thinking, and fourth, the author of this study speaks both of these languages.

6.1.2 Differences Between English and Japanese

To begin, we will compare English and Japanese because they significantly differ from each other, and these differences may trigger linguistic relativity. For example, English and Spanish are closer to each other than they are to Japanese. However, Spanish uses gender for nouns, but English does not, and this small difference might have a minor affect on perception (Boroditsky, 2003). Thus, this comparison will show linguistic differences that might influence perceptions differentially and where intuitions about these differences form the basis for empirical questions. For example, English is SVO while Japanese is SOV. Japanese speakers tend to omit the subjects of sentences, while English speakers do not. In English, plurality is common to almost all nouns whereas nouns generally have no inflected plural in Japanese.

These differences may provide starting points to test for linguistic relativity especially because bilingual speakers may perceive their subjective experience of relativity from differences like these. That is, the intuitions of bilingual speakers can form the basis of testable hypotheses. If differences between languages help facilitate differences in moral attention or thinking, the greater number of differences between languages also suggests the greater number of places to check for such effects.

6.1.3 Semantic Differences Between English and Japanese

Besides grammatical differences, perhaps a more promising area to research comes from semantic differences between English and Japanese. For example, Japanese has a large vocabulary for expressing register in honorific language and humble language. This might cause Japanese speakers to attend to levels of register more naturally than speakers of languages that do not tend to depict register so distinctively. Along these lines, it is interesting to note the degrees of register indicated by the numerous words for "I" and "You" in Japanese. The following chart lists a number of them, from most polite at the top, to least polite at the bottom.

Figure 6.1: You/I Pronouns in Japanese.

YOU	I
<i>otaku</i>	<i>uchi</i>
<i>sochira</i>	<i>kochira</i>
<i>anata</i>	<i>watakushi</i>
<i>kijo</i>	<i>watashi</i>
<i>kimi</i>	<i>ware</i>
<i>omae</i>	<i>washi</i>
<i>kisama</i>	<i>atashi</i>
<i>teme</i>	<i>boku</i>
	<i>ore</i>

For example, these many words for "you" might cause Japanese speakers to attend to levels of register more readily than speakers of Spanish who have just two words for "you" (*tú* and *usted*). Thus, we may ask: would Japanese speakers tend to look for finer YOU/I distinctions when learning another language even though the distinctions do not exist in it? Would such linguistic differences cause learners of Japanese to attend more to the status of speakers when speaking English? All languages can express respect and degrees of status. However, cultures differ significantly in how and how much they express register, and this seems like a good semantic space to check for linguistic relativity.

For example, Japanese speakers of English might say that English is good for being direct, but not so good for being polite. Interestingly, in conversation, Japanese speakers will tend to use a person's first name (or last name with the title *san*

afterwards) rather than the "you" pronoun. Perhaps using the name is less direct than using "you" and raises the question how usage of "you" in Japanese might affect how J/E bilinguals think about and use "you" in English. Hence, linguistic and semantic factors may cause J/E bilinguals to attend to politeness differently, tacitly priming them to see or express politeness where it does not exist in another language and culture. Nevertheless, these kinds of intuitions, anecdotes, and real linguistic differences need to generate empirical questions.

6.1.4 *The Intuitions of Bilingual Speakers*

Besides these cursory remarks about Japanese and English, another area that supports this study comes from the intuitions of J/E bilinguals. Boroditsky (2002) claims that bilinguals often intuit that language shapes their thinking and that they think differently in their different languages. Of course, we face the challenge to empirically verify these intuitions, but they are helpful, as Boroditsky claims, because they can inspire experiments. The following chart gives some anecdotal evidence from native English speakers who are bilingual in Japanese (E1J2) and native Japanese speakers who are bilingual in English (J1E2). These comments, which I have personally heard in over a decade of living in Japan, (unless otherwise referenced) may provide insight into how to test for the influence of language on thought for J/E bilinguals.

Figure 6.2: J/E Bilinguals Intuitions Related to Linguistic Relativity

E1J2	I find it much easier to be unequivocal in English than in Japanese.
E1J2	English speakers tend to be so hard, cold, and rational compared to Japanese speakers.
J1E2	Japanese is better for being polite in social communication than English.
E1J2	We (missionaries) find it hard to explain the English word "sin" in Japanese because the Japanese word <i>tsumi</i> means crime and does not convey the same meaning in English.
J1E2	Japanese people tend to respond more positively to relational interaction than logical interaction.
J1E2	I find it hard to express the English idea of personal independence and self-reliance in Japanese even though we have the term <i>jiritsu</i> for it.
J1E2	I find it hard to explain certain concepts in Japanese because of the lack of definite articles, plural words, or a clear word for that concept.
J1E2	Due to the ambiguity of Japanese grammar, their language cannot make exact sentences in past or future tenses... Moreover, without unambiguous sentences, it is difficult to think clearly and logically. (Terashima, 2003)

J1E2	Article Title: The Japanese Language Which Prefers Ambiguity versus The English Language Which Prefers Battle of Words (Inoue, 1998: 91)
------	--

These anecdotes demonstrate the difficulty of being bilingual and the difficulties that translators face. However, in this thesis, I will interpret them as anecdotal evidence for *linguistic relativity light*, pointing to the need for empirical research. Moreover, Condon highlights a number of bilingual intuitions that anecdotally support linguistic relativity and point to the need for empirical study.

There are at least 10 words in Japanese that might be equivalent to the English "I," and another 10 for "you" depending upon a particular relationship. Relative age, status, role, familiarity with each other, whether or not both are men or both women or man and woman together will all influence how Japanese talk. It is not just the pronouns that are truly relative; other features of speaking are also adjusted. This is why some Japanese say that they "feel more individualistic" when speaking English and why for no-nonsense talk English is a far better language than Japanese, at least from the point of view of many Japanese. "English is a perfect language for lawyers," said one Japanese friend, "but not for gently getting to know each other. English is like algebra -- hypothetical, impersonal, practical, modern, direct, whenever. It is easy to say 'yes' or 'no' English. That is what I like about it, it makes sense, is so modern." (Condon, 1984: 11)

When we think about linguistic relativity, these anecdotes are interesting. For example, though some J/E bilinguals allege that Japanese language prefers ambiguity, we would be incautious to conclude that Japanese language causes its speakers to prefer ambiguity while English causes its speakers prefer clarity. Japanese speakers can speak unequivocally, and English speakers can easily equivocate; moreover, clear counter-examples may exist that contradict the bilingual intuitions mentioned above. That is, one can simply say that Japanese speakers culturally prefer ambiguity -- not that the language actually causes them to prefer it.

Nevertheless, having stated this objection to a Japanese linguist, she replied that being unequivocal in Japanese causes one to speak incorrectly, (Yanagisawa, 2002, personal communication). Perhaps this means that being unequivocal in Japanese causes one to speak *in uncommon collocations*. However, if her claim is correct, then Japanese grammar or lexis tends to cause speakers to be more indirect and ambiguous than English. This is a far cry from linguistic determinism, but it shows

how language might influence differences in thinking and communication. Moreover, it validates research into how commonly collocated grammar and lexis might influence lesser or greater clarity in communication or how differential coverage of semantic space might influence thinking differently across cultures.

6.1.5 Empirical Evidences for Linguistic Relativity

Hence, these linguistic differences and intuitions support research in linguistic relativity -- comparing Japanese and English. This research does not come from a longing to save a lost cause in linguistics, or to prove the superiority of one language over another. On the contrary, we base this study on the real and clear intuitions and experiences of bilingual speakers and the significant semantic (or grammatical) differences between Japanese and English. Obviously, cultural values, instead of linguistic factors could be the only culprit here; however, if we can isolate even some linguistic influences from cultural ones, then we would have evidence, albeit subtle, for linguistic relativity.

At this point, it is important to briefly discuss some of the current evidence for linguistic relativity. Motluk summarizes the current state of affairs.

The general consensus is that while the experiments done by Lucy, Boroditsky and others may be intriguing, they are not compelling enough to shift the orthodox view that language does not have a strong bearing on thought or perception. (Motluk, 2002: 37)

Nevertheless, the consensus also is that the new approaches are more rigorous. Moreover, one main defense against linguistic relativity comes from language's apparent lack of influence on color perception. However, the new advocates of the theory claim that language may have its strongest influence in abstract spheres, such as "concepts of time, love, numbers and political ideas, where sensory information can't really help" (Motluk, 2002: 37). Incidentally, Belpaeme's research and computer simulations on color term acquisition caused him to conclude, "category acquisition coupled with language through cultural learning can indeed explain color sharing" (2002: 169). Moreover, experiments have also shown that an influence of language on [color] categorization is possible and even beneficial, thereby confirming the theoretical plausibility of the Sapir-Whorf thesis.

Nevertheless, leaving the color debate aside, for this discussion it is important to remember that key researchers in linguistic relativity claim that abstract and non-sensory spheres are potentially better for this research. This supports this study since it deals with abstract moral concepts that also tend to vary across cultures. Others might object that linguistic influence on thought implies that the language-effect on a people might be negative compared with the affect of a putatively better language. However, Boroditsky (2002) claims that her research reveals the flexibility and adaptability of language learners. That is, one's first language does not lock a person into a set way of thinking because if we teach people a new language, they also learn a new way of thinking. However, most essentially, the kind of influences language may have on thought are so subtle and nuanced that they do not produce the kind of differences where one could claim any kind of linguistic or cultural superiority.

For example, in one study, (Boroditsky, 2001) Mandarin and English speakers were compared in the ways they think about time. The English language tends to couch time in horizontal terms: looking *forward* to tomorrow, but falling *behind* the times. However, Mandarin speakers, besides having some *front* and *back* spatial metaphors for time, also use vertical metaphors about time. Hence, early events can be referred to with a morpheme for *up*, and later events with one for *down*. Boroditsky's study showed that Mandarin speakers tend to think vertically about time, even when using English. For example, (Boroditsky, 2002), Mandarin speakers confirmed more quickly that March comes earlier than April if they had just seen a vertical array of objects than if they had just seen a horizontal array, and interestingly the reverse was true for English speakers.

In addition, it appears that vertical time thinking in Mandarin-English bilinguals correlates to how old they were when they learned English, and English speakers taught to think about time vertically can develop the same vertical bias in priming experiments. Thus, Boroditsky concludes that language can indeed shape thought, that one's native language can influence habitual thought, and that people can develop new ways of thinking.

6.1.6 *Linking Linguistic Relativity and Moral Perception*

Up to now, we have focused on justifying study on how language might shape moral perception and thought. We have seen grammatical differences between Japanese and English and highlighted lexical differences where one bit of apparently full semantic space in Japanese appears lexically empty or lacking in English. For example, Japanese will say *itadakimasu* to express gratitude before eating, which literally means, "receive." Japanese students will ask how to say this in English, and when they learn that English speakers would say "Thank you" or "Let's eat," they feel dissatisfied because they want to say something different. We can say the same for the term *otsukaresama-deshita*, which means something like: "You must be tired," but is dynamically translated as "Many thanks," or "Thank you for your hard work." Hence, this represents another unsatisfying semantic space in English for Japanese speakers. Additionally, bilinguals have strong intuitions about how these linguistic differences between English and Japanese influence differences in thinking about politeness, directness, clarity, fuzziness, and rhetorical discourse.

To sum up this section, we need to remember our specific question on linguistic relativity. How do linguistic differences shape thought and perception, concerning categorization, identity, and moral development? Moreover, we need to remember the factors that justify this study: the intuitions of J/E bilinguals, suggestive evidence from new studies on linguistic relativity, and significant linguistic and value differences between Japanese and English speaking cultures. With all this in mind, we will look at the background and results of the Moral Perception Experiment.

6.2 **Methodology of the Moral Perception/Dilemma Experiment**

In total, 41 J/E bilinguals participated in this study (14 in a preliminary study and 27 in a secondary study). They varied in their bilingual abilities, but they all explicitly declared enough bilingual ability to complete the survey, except for two participants who indicated that their Japanese ability was not sufficient. They were not included in the data. The subjects for the secondary experiment were recruited from Internet listservs where English-Japanese bilinguals might be found, such as a

group of teachers of Japanese in North America, or from The Linguist List. These participants did the survey over the Internet on a web form.

6.2.1 *Preparation and Piloting*

The survey aimed to answer this main question: Attempting to minimize all variables except language, will J/E bilinguals respond differently to equivalent translations of moral dilemmas mainly because of differences in language?

Initially, the instrument was piloted in English to refine the questions. Then the original survey was translated into Japanese from English, and after that a separate translator translated the survey back into English from Japanese. We compared the back-translation with the English original, and based on discrepancies between the original and the back-translation, we revised the Japanese translation to enhance its equivalence to the original English. Hence, we made sure the surveys mirrored each other as faithfully as possible. Subsequently, the preliminary study presented the mirror translations of the surveys in English and Japanese to 25 J/E bilinguals, and 14 of them responded. The survey consisted of 15 moral dilemmas, like this one here. Note that all of the questions in English and Japanese can be found in Appendix I.

You find 10,000 yen (\$100 US) on the street, and you are alone and no one can see you. Would you keep the money (A) or turn it into a police station (B)?	<input type="checkbox"/> A	<input type="checkbox"/> B
--	-------------------------------	-------------------------------

The preliminary survey was first given in Japanese, during the week of February 18-22, 2002, and the corresponding English survey was administered during the week of March 4-8, 2002. The time separation between the surveys aimed to decrease the influence of the first survey over the second survey, (Ervin-Tripp, 2001: personal communication). Subjects were reminded that the survey questions were moral dilemmas that happened to real people, and though their identity remained anonymous, surveys were numbered so that the researcher could compare answers for individuals between surveys. For Japanese native speakers (all but one of the subjects), we considered the following sufficient to include them in the survey: at least one year of study abroad, a completion of a university degree in an English speaking country, or a significantly high TOEIC score. The one native

speaker of English possessed a superior level of Japanese and easily responded to the survey.

6.2.2 Results of the Preliminary/Pilot Survey

Of the 14 subjects who completed both the Japanese and English sides of the survey, only one subject returned the exact same answers on both surveys. Moreover, though individuals gave the same answers on a majority of the questions on both surveys, more importantly, 13 of the 14 gave at least one different answer for the same question. The average discrepancy for individuals was 2.3 answers per person, and the total spread set out as follows:

No. of Subjects	Ratio	Incongruity %
1 Subject	15/15	0
3 Subjects	14/15	6.67%
3 Subjects	13/15	13.33%
5 Subjects	12/15	20.00%
1 Subject	10/15	33.33%
Average	2.3/15.00	15.33%

Though patterns of difference were not overwhelming, they were significant enough to justify further study. For example, 38% or 5 of the 13 subjects expressed incongruity in 20% or 3 of their answers. If these participants truly understood the questions in both languages, and if we have isolated language as the only *significant* variable, then this result supports the need for further research.

However, a number of issues arise from these preliminary data. First, the subjects were sufficiently proficient in both languages to insure that lack of comprehension did not cause the incongruities, but proficiency would need to be more carefully checked. This would increase confidence that differences were not based on misunderstanding the dilemmas. Second, when a higher number of subjects showed incongruity between Japanese and English answers for the same questions, they sometimes gave more conflicting responses. That is, for some questions, answers went in both directions, and no pattern emerged. Obviously, answers going in both directions in the pilot study make the data hard to interpret, but these answers could indicate that Japanese answer vaguely in Japanese and more explicitly in English. Perhaps this interpretation is simply guided by the hypothesis.

However, because some subjects gave answers that went *consistently in one direction on some questions* (for example, preferring to confront in English and not confront in Japanese), we judged it worthwhile to administer the survey a second time in order to get more data. In preparation for this, a back-translator and this researcher carefully checked the survey questions again and refined the accuracy of the mirrored surveys.

6.2.3 *Methodology and Results of the Secondary Survey*

As already mentioned, subjects took the secondary survey online. Because of the flexibility of the Internet, the experiment did not have an official beginning and ending, so the only key time factor was an individualized delay before the subjects did the second language part. Besides being Internet-based, another aspect that was changed from the preliminary survey was that approximately half of the subjects did the English questionnaire first, and half of the subjects did the Japanese one first. This insured that different answers did not come from the order in which subjects responded to the dilemmas.

The subjects ranged from 18 and over 40 years of age, averaging about 30. Nine of the subjects were male, and eighteen were female; nine had attained graduate level education (Master's Degree), nine post graduate level (PhD), eight undergraduate level, and one had graduated from high school. Reading ability was a key factor, and the survey checked reading with this question: "What is the level of your reading ability?" Out of the twenty-nine subjects who completed both halves of the study, two indicated a junior high reading level in either Japanese or English, and therefore their responses were deleted from the data. For the other participants, in Japanese, 21 indicated they had graduate level ability, 5 indicated undergraduate ability, and 1 indicated high school reading ability. For English, 25 indicated they had graduate level reading ability, and two indicated they had undergraduate ability.

6.2.4 *Discussion and Results of the Moral Dilemma Survey*

We hypothesized that the two languages will differentially affect moral thinking -- at least in subtle ways. Due to the challenging nature of the experiment, we need to

consider this an *exploratory experiment*. Nevertheless, we addressed foreseeable problems, with careful piloting, back-translations, checking the bilingual ability of the subjects, protecting their anonymity, putting a delay between both surveys, and having some subjects do English first and others do Japanese first.

The following chart summarizes questions where results were most suggestive. The numbers in the chart below indicate the sum total of movement in one direction. For example, on Question #3, 5 subjects preferred answer A in English and B in Japanese, and only one subject preferred A in Japanese and B in English; hence, the sum total of the "E-movement" is +4 for English.

Question	E-Move	Percent
3. Your co-worker always comes late to work and dresses poorly. You know the boss is unhappy and is considering firing your colleague. Would you warn her (A) or avoid confrontation (B)?	+4A	14.81%
12. Your friend invites you to a party, but you do not want to go. Would you tell him directly that you are not interested in going (A), or would you think of some polite or indirect excuse for not going (B)?	+4A	14.81%

For questions 3 and 12, a potentially interesting pattern emerges: both questions deal with directness and indirectness. 15% of respondents indicated that they preferred speaking directly in English than in Japanese. Specifically, 4 of 27 subjects in the secondary study indicated that they would warn a colleague *in English* about her behavior; however, they would avoid the confrontation in Japanese. This question partly concerns manners (a kind of mini-morality); however, it also carries moral implications, such as one's responsibility (or lack thereof) for protecting the livelihood of the lazy co-worker and the co-worker's responsibility to do her work. Confrontation can also carry moral implications because sometimes we find ourselves in situations where ethical and moral issues call us to confront people who are breaking the rules. If your language even slightly causes you to avoid that confrontation, then linguistic factors influence the moral outcome.

Regarding question number 12, 15% of the respondents indicated that they would directly turn down an invitation in English but politely excuse themselves in

Japanese. Though this dilemma deals even more in manners than question 3, the result still hints at a moral dilemma. For example, though this invitation is benign, human beings often tempt others into moral dilemmas. Concerning the Japanese tendency to value harmony and the inclination to esteem indirectness, cultural influences obviously play a dynamic role, and these cultural pressures could easily override language.

For example, a Japanese bilingual with a propensity to value harmony, might make a polite and indirect excuse *in English* to turn down an invitation. Any serious and careful discussion of language, culture, and linguistic relativism must admit the real influence of culture on thought -- separated from the influence of language on thought. The challenge will always remain to tease apart the linguistic influences from the cultural ones. Therefore, perhaps one of the best forms of evidence in favor of linguistic relativism regarding the harmony/directness distinction would surface in the probabilistic tendency of bilinguals to prefer English to be direct and Japanese to maintain harmony. The result for question 12 suggestively highlights this tendency.

Question 12 is also interesting because it resembles Question 3. Both these questions deal with directness, indirectness, clarity, and vagueness. The simple result appears that bilingual speakers of Japanese and English might prefer to respond less directly in Japanese and more directly in English. Moreover, if Japanese L1 speakers follow this pattern, and if English L1 speakers do not, this does not discount this result. This would simply mean that native speakers of Japanese prefer directness when speaking English and indirectness when speaking Japanese, still indicating some kind of differential linguistic influence on thinking.

Interestingly, the answers were split 50/50 on question 12. That is, one would expect only the Japanese L1 speakers to prefer directness in English and ambiguity in Japanese. However, for question 12, two Japanese L1 subjects moved to English (E-moved), preferring directness in English, and from the others, one was a native English speaker, and the other Italian. For question 3, three respondents who E-moved were Japanese L1 subjects and the other two were English L1 subjects. This might indicate that regardless of the L1, J/E bilinguals tend to prefer confrontation

and directness when speaking English and harmony and ambiguity when speaking Japanese. *Thus, to know English (regardless of your first language) is to know a way to be direct, and to know Japanese thus means to know a way to be indirect.* However, the most interesting result relates to differences in answers between males and females. In the secondary survey, for question #3, 5 of 18 females preferred to avoid confrontation in Japanese but have it in English, but none of the 9 males moved on this question. This suggests that female J/E bilinguals might tend to prefer harmony in Japanese but directness in English.

6.2.5 Further Statistical Interpretation of Questions 3 and 12 (Secondary Survey)

Japanese first or English first. 14 subjects took the survey first in Japanese, and 13 took it first in English. For all 15 questions, results showed that Japanese first or English first had no affect on the pattern of the answers (Fisher's exact test, $p = .098-1.00$). For question 3 and 12, the subject's native language did not influence the outcome in any statistically significant way. For question 12, there were two answer patterns: no change and from B to A. We compared the answer patterns of question 12 between subjects whose first language was Japanese and those whose first language is other than Japanese. Using Fisher's exact test, the p-value for question 12 was .613, indicating there was no statistically significant relationship between the answer movement of question 12 and the subjects' first language. For question 3, there were three answer patterns: no change, from B to A, and from A to B. There was no statistically significant relationship between the answer pattern of question 3 and the subjects' first language (Chi-square test, $p = .734$). This shows that the answer patterns were almost the same for Japanese as first language and other language as first language. In a word, results did not significantly change whether subjects took the survey in Japanese first or English first, or whether their first language was Japanese or another language.

Regarding gender factors, the following data for question #3 are most suggestive. In the secondary survey for question #3, we see that 6 of 18 females chose directness in English and avoided ambiguity in Japanese (McNemar test, $p = .145$), and none of the males moved. Although a statistically significant relationship between the answer patterns of question #3 and gender was not obtained,

obviously, these results are most suggestive, and we may attribute this lack of significance to a small sample size. Thus, this result at least suggests that with more powerful data, we might obtain a statistically significant result regarding the relationship between gender and the answer patterns for question 3. Question 3 deals with confrontation, and our results raise the question: would women tend to avoid confrontation in Japanese while attempting it in English? These results are significant enough to justify a larger study looking at how gender and language might influence behavior related to confrontation, harmony, and directness.

These results do not confirm our hypothesis; however, they encourage us to further research how language may influence thinking and behaving. Moreover, our results highlight the need to research how gender and language might affect harmony and directness. J/E bilinguals confirm by their intuitions that linguistic differences affect the harmony/directness factor, and we know that gender differences influence behavior also. Personally, as an American, who has lived over a decade in Japan, I know that I culturally prefer direct communication, and this can cause me to speak too directly in Japanese. However, the question is whether I will at times tend to speak more indirectly (and harmoniously) not simply because of the influence of Japanese culture on me, but rather because of the influence of Japanese language.

6.3 Additional Reflections on Linguistic Relativity and Morality

If Japanese language can influence a bilingual speaker towards harmonious speech behavior, it could also affect one's thinking in Japanese -- since Japanese common collocations tend to express more harmonious and indirect speech. On the other hand, Japanese could influence a person to prefer harmonious speech in English. It is hard enough to tease apart the influences of language and culture. However, let us assume that Japanese language gives me some nuances that facilitate harmony or modesty, and English does not as readily give me these nuances. Then a situation gives me the opportunity to employ those nuances where I also have a choice to use English, and I choose Japanese, to consciously facilitate harmony. That is, I may use Japanese at this point because it gives me nuances that English

does not; thus, these available nuances cause me to speak in Japanese, not in English, and this would show an indirect influence of language on my thinking.

Thus, intuitively, I would claim that *the linguistic resources available to me represent an indirect influence on my thinking*. Metaphorically, these linguistic spaces available to me are rooms in a large house. The word *itadakimasu* represents a non-room or underdetermined-space in English because we cannot translate it directly. The verb means, "receive" but is often translated as "Thank you." This is a dissatisfying translation because it represents a *thought-room* that I cannot directly enter in the English part of my thought house. It is a non-space or underdetermined semantic space because no simple translation exists in English.

Therefore, I can only completely enter into this thought space in Japanese, or through an unwieldy English translation that does not represent the original meaning. That is, even though I can roughly translate *itadakimasu* into English, the translation does not fit the common linguistic and situational collocations in English language and cultures. If a speaker cannot easily enter a particular semantic space in one language, but may freely enter it in another language because of the different linguistic resources available, then this represents a valid case for language subtly shaping and influencing thought.

These issues become even more complex because we find it difficult to distinguish between language and thought. Though we may think acoustically (as in music) or visually (as in painting or photography), I find it hard to do any general thinking without language. Whatever the difference between language and thought, words may be the keys to *thought-rooms* in our thought house. If a word exists in Japanese in a highly economic form and not so in English, this word could also be a key to a *thought-room* in Japanese, a room that English does not completely build. We still might be able to build this *thought-room* in the English wing of the bilingual's thought house, but we might need to coin a key word or words that would help us enter that semantic space.

For example, Japanese language has the phrase *enryo-no-katamari*, which refers to the last amount of food or drink, leftover in a dish or a pitcher that people avoid

using up. *katamari* is translated as "lump" or "mass," and *enryo* as "reserve, diffidence," or "modesty." Hence, we can translate the phrase as "the mass of modesty," "the remains of reserve," or "the diffidence remainder." However, we cannot usefully and economically express these phrases in English without such coinage *and* explanation, even though English speakers may experience this phenomenon. Translation is possible; however, the phrase in Japanese represents a *thought-room* that is a non-space (or under-determined space) in English. Bilinguals open this *thought-room* with a unique keyword (unless they coin a phrase and intentionally explain it in English). Therefore, at points like this, language may differentially influence thought because the *thought-room* and therefore the thought are only fully or easily available in one of the two languages.

This raises the question of whether there might be more than referential spaces delimited by thought-rooms. Could moral thought-rooms exist, semantic spaces related to moral values that might find fuller expression in one language and culture than another? If such value-laden linguistic differences exist, culture would likely play a significant role in creating them. However, this would not necessarily discount the potential influence of language on thought. Even if culture carves out all the value differences between two cultures, members of these cultures would still understand and *think* of these differences primarily through language.

Related to this, when we think about a Japanese value that English underdetermines, how else can we think about this value, except primarily through the Japanese language? This thinker will need to come up with a term like "modesty-mass" in order to explain the Japanese value in English. The Japanese language did not create this value, but Japanese language enables us to symbolize it, and interestingly, since English underdetermines it, we cannot easily communicate it in English. For the bilingual speaker, the influence of language on thought crystallizes at the point where we can clearly express a value in one language but can only under-express it in another language. In this way, language only shapes thought to the extent that it provides well-carved channels for our thoughts. If so, when bilinguals express a particular value, they naturally prefer to

express it in the language that most economically symbolizes it -- where that value is most fully linguistically determined, and not under-determined.

This is not linguistic determinism, which implies we cannot think of or understand concepts for which our language has no words. On the other hand, this suggests that we prefer to express thoughts in the language in which they were created. Moreover, we will be at a loss for words (and their thoughts) when we try to express them in a second language that under-determines these thoughts. Hence, future research on language-shaping-moral-thought might benefit from focusing on morally-value-laden words that bilinguals can most concisely express in one language and not another.

The survey outlined in this chapter represents one experimental approach to ask how language might influence thought, especially in the moral realm. Specifically, can linguistic differences between English and Japanese influence J/E bilinguals to prefer different solutions to moral dilemmas depending on the language they respond in? This is a bold question, and therefore we can understand why we did not obtain a significant result; however, this is actually quite reassuring. That is, if a high number of subjects gave different answers on a high number of dilemmas, then this would reflect a world out of kilter where people do very different things depending on the language they thought in. Bilinguals would be strange -- behaving very differently depending on the language they were thinking in, and language would over-determine and perhaps warp our understanding of the real world.

Incidentally, the linguistic and paralinguistic behavior of bilinguals might make them vulnerable to the charge of hypocrisy. Anecdotally, as a speaker of Japanese, Spanish, and English, I recall a meeting where I greeted and spoke to Spanish speakers in Spanish, and Japanese speakers in Japanese. One observer called me a hypocrite because I behaved differently in each conversation, bowing like a Japanese one minute, and gesticulating like a Latino in another. This matches the kinesic differences that are often observed in bilinguals:

Speeches given by New York's colorful mayor, Fiorello LaGuardia, who spoke fluent Italian, Yiddish, and American English, illustrate how closely kinesic activity is linked to culture. An observer familiar with the three cultures could watch LaGuardia on a newsreel film without a sound track and tell readily which of the three languages he was speaking. There seems to be a subtle shift of kinesic gears when a fluent speaker slips from one language to another. (Morain, 1987:122)

Nevertheless, if culture or language can affect the kinesic behavior of bilinguals, this is still somewhat insignificant compared to the idea that linguistic differences affect their moral choices. Moreover, bilingual intuition predicts that language influences thought in very subtle ways, as the results suggest for questions #3 or #12. For example, when turning down an invitation, language might influence *the way* one solves a dilemma, as opposed to determining *the result*. Perhaps the result is the same -- the speaker turns down the invitation, but in different ways depending on the language. For example, a J/E bilingual may turn down an invitation more subtly and indirectly than she would in English because of the linguistic and semantic resources that Japanese provides.

The cause could be simple. In English, J/E bilinguals may consider the common collocations of direct speech to be comparatively polite, but in Japanese, bilinguals may consider direct translations of those common collocations to be impolite. That is, compared to English, Japanese may not have as many common collocations in direct speech that speakers consider polite. These differences in collocations could influence speakers toward different forms of verbal behavior -- and thought that motivates behavior. If one avoids a confrontation in Japanese but has it in English, the result would probably be different. However, based on our data, differences in Japanese and English might only influence a different result (as opposed to manner) when the situation forces a choice between harmony and directness.

6.4 Three Summations

Three conclusions emerge from our data. First, language may subtly influence moral thinking -- more regarding *the way* bilinguals solve certain moral problems, rather than *the result* they choose. Secondly, we need to pay attention to the influences of both language and gender. Third, future research needs to focus on

these factors. This survey covered many moral issues, and at this stage this was necessary. This gives background and sets out the issues and problems. Therefore, future studies could focus on one moral issue, include the gender factor, and deal with the ways bilinguals solve moral problems as well as the results they choose. Moreover, besides using moral dilemmas, we should find other ways to test for how language might influence moral thinking. Hence, the following section briefly suggests other means to research how language might influence moral perception.

6.4.1 *The Fitness of Concepts*

For future research, we might employ the Semantic Differential Technique (SDT) (Osgood and Tannenbaum, 1957) For example, subjects could evaluate morally laden words, using the SDT categories: Evaluation (good/bad), Potency (strong/weak), and Activity (active/passive). Subjects would rank best-possible translations of moral terms in the following way.

Good	3	2	1	Japanese Word X	1	2	3	Bad
Good	3	2	1	English Word X	1	2	3	Bad

This technique might show us how J/E bilinguals differentially evaluate best translation pairs, and if a significant number of bilinguals evaluate these pairs differently, then this could show a linguistic influence on thought. That is, if I think Word X is stronger or better in Japanese than its best translation pair in English, then my semantic resources cause me to think differently about words that I use to convey the same meanings in both languages.

6.4.2 *Translation and Underdetermined Semantic Spaces*

This possible experiment raises an interesting discussion that may help us better understand the robustness or fitness of ideas in cross-cultural translation. That is, if a concept is hard to translate, or receives a lower evaluation in Japanese than English, then this concept may experience lower fitness in Japanese culture also. We translate the word "sin" as *tsumi* in Japanese. The word "sin" obviously does not exist in Japanese because it is an English word, but more importantly, *tsumi* (as a translation pair) does not represent the same concepts as the English word "sin." The words have different nuances. If no Japanese word accurately represents the

concept of sin, then does this underdetermined concept in the Japanese cultural context mean that it will experience a lesser degree of fitness in Japanese culture than western English speaking culture? If yes, then this might breathe life into the currently floundering and unpersuasive field of memetics, which is the topic of the next chapter. Thus far, research has not produced good evidence for the existence of real memes. However, if memes do exist, the idea of memetic fitness, (conceptual fitness) suggests that some memes might experience differential fitness depending on the cultural environment they reside in.

The issue of under-determined semantic spaces suggests that when people misconstrue difficult to translate words this may constitute a form of linguistic relativity. Hence, even though the two words, "sin" and *tsumi* are best translation pairs, their meanings differ. Though a true bilingual should have the linguistic and cultural competence to understand these words in their cultural contexts, he initially may only partly understand the different nuances of this translation pair. *Tsumi* connotes something closer to the idea of a crime in Japanese, and "sin" refers to a violation of a religious, ethical, or moral standard in English. Most importantly, *tsumi* does not generally imply a transgression against God's law as the word "sin" does in English. The words do not share the same cultural backgrounds.

Thus, the different meanings could cause a developing bilingual to think the Japanese cultural concept for "crime" when responding to the word "sin" in English. Thus, until the learner further develops his linguistic competence, his first language may cause him to translate and infer in Japanese different nuances for the word "sin" than its meanings in English. Because of differentially developed semantic spaces, a particular word can influence developing bilinguals to think of nuances that this word does not carry. My English word "sin" connotes the cultural concepts connected with it, but when I use *tsumi* to say "sin" in Japanese, *tsumi* can connote a different set of concepts. In short, these different words -- though best translation pairs -- influence listeners to think different thoughts.

This happens because the pair actually signifies different meanings, but the learner thinks they are the same because her L1 word causes her to think in this way.

Though culture influences her understanding, depending on her background, language may still trigger the same nuance for both words when they mean different things. Clear and simple translation word pairs are very common (e.g., for physical objects), and this may cause learners to expect this clear linkage in cases where it does not exist. The culturally naïve learner sees the linguistic translation between cultures as something that fully determines an equivalent meaning; however, the language actually under-determines the meaning in many cases. Thus, when language under-determines meanings across cultures, linguistic relativity may occur.

In this scenario, the translation-pair serves as a link between the two cultures. When the best available translation pair under-determines the link, that is, where each word cannot express its partner's meaning without a cultural explanation, then this may cause linguistic relativity. In this case, language influences thought because it underdetermines the link between cultures. Interestingly, here linguistic influence over thought does not come from the power of language, but rather from a gap in the language. That is, the translated word-pair expresses a meaning-pair that for difficult-to-translate words stands as a vague meaning-pair, or under-determined meaning-pair. Hence, when we symbolize mutually under-determined concepts across cultures, the symbols cannot accurately represent the cultural concepts we intend them to convey. Thus, language will shape thought across cultures in ways that it might not within one culture because translations can sometimes under-determine their cross-cultural meanings. The best translation-pair still represents an inaccurate translation, so it can influence thought inaccurately. In short, the best-available translation pairs can influence us to think different thoughts than the translator intended. These translations under-determine meaning causing a form of linguistic relativity.

6.4.3 *Linguistic Relativity and Ethnography*

This problem relates to a central issue of ethnographic research *called translation competence*, which results when would-be ethnographers interpret meanings from one culture in terms of their own culture. Ethnography scholar, James P. Spradley comments:

Language not only functions as a means of communication, it also functions to create and express a cultural reality. When ethnographers do not learn the language, but instead depend on interpreters, they have great difficulty learning how natives think, how they perceive the world, and what assumptions they make about human experience. The barrier to learning their particular frame of reference, their cultural reality, has not been removed. The more an informant translates for your convenience, the more that informant's culture reality becomes distorted. (Spradley, 1979: 20)

Spradley may overstate the power of language to create a cultural reality; nevertheless, in ethnography when we depend on interpreters and translation, this relates to the issue of under-determined meaning-pairs mentioned above.

Translation competence describes the ability to translate a meaning in one culture into a form readily understood in another culture; however, the original cultural meaning may not exist in the target culture. Thus, meaning-form-X comes to mind in the target culture when meaning-form-Y is the actual meaning in the origin culture. For example, one could translate *sukiyaki* as "stew" or *sushi* as "raw fish" in English, but these translations do not perfectly convey the Japanese meanings, and this is probably why we use these Japanese words in English.

In these cases, translation competence becomes *translation incompetence* because it causes us to mistranslate and misunderstand intended meanings. Nonetheless, it produces a form of linguistic relativity, where simple word-pair translations impede accurate thinking about cultural concepts, especially for word-pairs whose meanings are cross-culturally under-represented. In fact, to avoid translation competence, ethnographers often rely on the original words to convey concepts on their own cultural terms. This conventionalized reliance on original words in ethnographic research provides good anecdotal support that semantically under-determined translation-pairs cause a *light form* of linguistic relativism.

Moreover, relying on the original words creates problems for translation. If we need to rely on original words, then translators will face serious difficulties translating some concepts between cultures. For example, the word "sin" will connote guilt and violation of moral law in English, but translated as *tsumi*, it will connote slightly different meanings of shame and violation of group harmony in Japanese. This is a problem for translators: however, our focus concerns the

theoretical issue of how languages influence thought, especially moral thought. Though some grammatical features may shape thinking (Boroditsky, 2002), the word-pair issue relates more to the nature of language than specific examples of syntax or metaphor type. Words in Language X are culturally bound to the culture(s) that use Language X. When we translate a word from Language X to Language Y, we sometimes cannot convey Language X meanings in Language Y. The meanings are culturally determined, and translation underdetermines them. Thus, as literal word-pair translation underdetermines meaning, it also shapes thinking. Thus, language in translation can misshape thinking, away from the original meaning of the word-concept.

6.5 Conclusion

This chapter focused on the relationship between language and thought and how linguistic differences might influence moral thinking and choice. The results of our experiment suggested the need for further research on how language might influence some forms of morally laden interaction. It will be very difficult to disentangle the interdependent influences of language, thought, and culture. However, this experiment may have hinted at a way, for example, for us to demonstrate that J/E bilingual females may tend to avoid direct confrontation in Japanese and not avoid it in English. Future experiments will need to focus more specifically on these factors. Moreover, we looked at how language in literal translation appears to underdetermine meaning and misshape thought away from original meanings. This may be especially true for moral concepts that can differ exceptionally across cultures. Along these lines, this discussion leads to an interesting question concerning the fitness of concepts (or memes) across cultures.

This idea of conceptual or memetic fitness concerns the subject of the next chapter, and I will introduce it here because it relates to linguistic relativity. Memetics is supposed to be a way to study the fitness of concepts or memes; however, this approach has not yet offered any novel results, and many scholars have understandably given up on it. See Aunger's edited volume where leading scholars critique the state of memetics (Aunger, 2000). However, if we look at the meme

idea cross-culturally, considering how some word pair translations underdetermine meaning, this might set up an interesting memetic question. If Idea A connotes nuances in Language X and the cultures of Language X, but semantic resources underdetermine those nuances in Language Y, then will this situation predict the fitness of Idea X in Language/Culture Y? For example, if *tsumi* as a translation for "sin" underdetermines the meanings in Japanese, then will this predict that these English based-nuances will experience low conceptual fitness in Japan?

Moreover, this question relates to language and morality in two ways. First, we may find some moral concepts hard to translate because some translation pairs underdetermine meanings. As already mentioned, though speakers of English may regard the concept of "sin" as a violation of moral law connected with guilt, speakers of Japanese understand its translation-pair, *tsumi*, as a crime against society that brings shame. Second, though culture plays a role, language also influences the way people think about these concepts, and this may affect the fitness of the western concept of sin in Japanese culture.

Thus, the problem is linguistic as well as cultural, and in order to diminish translation competence, we may need to create new words or follow the ethnographic method so that we can avoid underdetermining meanings. However, if we do this, we will want to know if the original words like "forgiveness, sin, or guilt" will experience conceptual fitness as translated words in Japanese because of their different nuances in English. In short, this highlights issues for future research, concerning how language may shape moral thinking and how language might underdetermine moral thinking across cultures. This sets out the topic of the following chapter, concerning *the fitness* of morally laden concepts across cultures, making an interesting connection with the focus of this chapter, which has emphasized how linguistic differences also might influence differences in moral thinking across cultures.

CHAPTER 7

LANGUAGE, EVOLUTION, MORALITY, AND MEMES

7.1 Language, Evolution, Morality and Cultural Transmission Theory

Up to now, this thesis has investigated numerous ways language relates to evolution and morality, including exploring how language differences might influence moral perception and categorization. In this chapter, we look at how this relates to memetics -- a Darwinian approach to cultural transmission that is controversial, mainly because it has yet to deliver results. Memetics may turn out to be a failed enterprise, but it is still too early to bury it in the graveyard of ideas. Some reputable scholars have advocated memetics (Dawkins, 1976; Durham, 1991; Dennett, 1995; Hull, 2000; Aunger, 2002; Balkin, 2003), and if we can refine theory and methods, memetics still might help us see how language interacts with morality. Thus, this chapter will first analyze and criticize memetics and then define a limited version of it that might help us better understand why "moral memes" might spread and survive.

Despite its problems, we may still find a few reasons to apply memetics to the study of culture and morality. First, though some memeticists have earned disfavor for the approach by colorful or exaggerated posturing (Blackmore, 1999) or prematurely calling it the new science of culture (Lynch, 1996) one memetic idea may still hold promise. An *independent replicator dynamic* may exist for concepts, as if some memes have a life of their own. An independent replicator dynamic implies that memes may replicate -- to a degree -- independent of human intention, or at the least, factors may exist that increase the reproductive success of some concepts.

Second, this thesis has argued that language fundamentally enables us to construct and maintain a moral code. Moreover, humans transmit moral codes from generation to generation; moral codes evolve over time, and moral codes may experience degrees of fitness due to their structure; therefore, a memetic-type approach for understanding morality might prove helpful. Third, as previously discussed, some moral codes may actually contribute to the higher fitness of one group in comparison with another group. If so, the moral code that contributes to the fitness of its group will -- as a code -- experience a greater fitness than a moral code that contributes less fitness to its group. With these things in mind, we will survey the broader terrain of memetics before dealing specifically with how it supports this thesis -- that language forms a significant foundation for the origin, grounding, and variable fitness of human moral codes.

As scholars have applied Darwinian approaches to an increasing number of academic disciplines, (See Hull, 2000) some have proclaimed this the Age of Universal Darwinism. Riding this trend, for the study of culture and mind, memetics employs the concept of memes as cultural and psychological replicators analogous to genes -- the biological replicators. We can consider memetics a form of evolutionary social psychology that reduces culture to bits and replicators. Nevertheless, before we can take it as seriously as biological evolution, memetics still faces serious theoretical and empirical dilemmas. These technical problems, some of which are discussed below, may appear insignificant at first. However, concerning many disciplines, including linguistics, anthropology, sociology, psychology, theology, and ethics, memeticists have made sweeping claims that touch on the origin of language, the meaning of personhood, and the grounding of epistemology and morality. See, for example, (Dawkins, 1976; Lynch, 1996; Blackmore, 1999; Brodie, 2004).

One problem for memetics comes from the strong and intrusive rhetoric of some memeticists. For example, the following quotations show the tenor of the debate and how some scholars perceive memetics.

The fallacy of Dennett's argument [strongly emphasizing the power of natural selection as opposed to Gould's stress on pluralism of processes in evolution]

undermines his other imperialist hope that the universal acid of natural selection might reduce human cultural change to the Darwinian algorithm as well. Dennett, following Dawkins, tries to identify human thoughts and actions as 'memes,' thus viewing them as units that are subject to a form of selection analogous to natural selection of genes. Cultural change, working by memetic selection, then becomes as algorithmic as biological change... thus uniting the evolution of organisms and thoughts under a single ultra-Darwinian rubric. (Gould, 2001: 97, comment in brackets mine)

Whatever we invent in our lifetimes, we can pass on to our children by writing and teaching. Evolutionists have long understood that Darwinism cannot operate effectively in systems of Lamarckian inheritance -- for Lamarckian change has such a clear direction and permits evolution to proceed so rapidly that the much slower process of natural selection shrinks to insignificance before the Lamarckian juggernaut. (Gould, 2001: 98)

Gould makes two important claims: (1) cultural change may be Lamarckian, or non-Darwinian, and thus (2) applying memetics to culture may misappropriate Darwinian theory. In fact, the picture may be somewhat more complex. Hodgson (2001) suggests that basic Darwinian principles of variation, inheritance, and selection can operate at the cultural level, and Lamarckian inheritance of acquired characteristics may operate on top of these Darwinian principles at the cultural level as well. If this is the case, then perhaps memetics can help us begin to clearly differentiate between these two kinds of cultural processes. Nevertheless, to date memetics has not empirically demonstrated the existence of either of these two kinds of inheritance occurring at the cultural level. This leaves us with the distinct possibility that memetics may not be able to help us find either Lamarckian or Darwinian processes of cultural change.

In addition, philosopher Mary Midgley strongly criticizes memetics as follows.

If memes really correspond to Dawkinsian genes they must indeed be fixed units -- hidden, unchanging causes of the changing items that appear round us in the world. But all the examples we are given correspond to phenotypes, [not genotypes]. They are the apparent items themselves. Moreover, most of the concepts mentioned [as examples of memes] cannot possibly be treated as unchanging or even as moderately solid. Such customs and ways of thinking are organic parts of human life, constantly growing, developing, changing, and sometimes decaying like every other living thing. Much of this change, too, is due to our own action, to our deliberately working to change them (Midgley, 2001: 76-77)

[A meme's] success is then due to its own reproductive strategy -- presumably produced by a mutation -- not to any fact about the people concerned. We need not look at those people. We need not relate the meme to these people's intentions. We certainly need not look at human psychology generally and look into our own hearts to see what we might learn there about such conduct. We simply place the whole causation outside human choice, thus avoiding that overestimation of our own powers which so disturbs Dennett. This effort to avoid pride would of course land us in a quite unworkable kind of fatalism (Midgley, 2001: 82-83)

Midgley raises numerous serious issues: the phenotypic aspect of memes, the issue of causation, and the factor of human intention, and this reveals just a few of the conceptual and empirical problems facing memetics.

To be fair to Dawkins, he develops and improves his definition of memes in *The Extended Phenotype* (1982). He admits that at first he did not sufficiently distinguish between memes as replicators and memes as phenotypes, that is, between memetic instructions, and memetic products. To clarify, Dawkins refers to Cloak's (1975) ideas of "i-culture" (instructions that people carry in their brains) and "m-culture" (material culture, technologies, and verbal ideologies). This indicates that at least memeticists have made progress regarding the replicator and phenotype distinction.

Nevertheless, though memetics has been around for over 25 years and attained some fame and infamy, the academic community has not provided a strong consensus regarding its *status as science* because of the contentious issues mentioned above. That is, though its most vocal proponents claim memetics is science, many critics disagree. Moreover, we have yet to see what memetics can contribute to our knowledge of culture and mind, two subjects that it is supposed to illuminate. Aunger (2000) says that scholarly opinion regarding memetics ranges from enthusiasm to disdain. However, perhaps we can disdain the negative aspects of memetics and enthuse over its potential.

Therefore, after providing further background on the negative aspects of memetics, I will focus on three major problem areas and two areas of potential. The problem areas include how memetics is: (1) ill-defined, (2) misguided and misinformed, (3) unhelpful and unscientific. The potential forms of memetics are (1)

"neuromemetics" (2002) and (2) a proposal that I will call lingua-memetics. The following section strongly criticizes memetics, but the goal is positive -- to find ways that memetics might help us better understand language, culture, and morality.

7.2 Basic Background: Where are the Real Memes?

For Dawkins (1976), who coined the current term, memes are units of culture that self-replicate with a life of their own. Blackmore (2000: 53) says, "Memes are stories, songs, habits, skills, [and] inventions." Durham (1991) defines a meme as a "cultural unit" where information actually affects behavior and transmits itself in a duality or multiplexity of means; and takes shape in discrete bodies and gets transmitted in coherent units. These definitions fit within a mentalist school of memetics, which beside the "meme as gene" school includes the "meme as germ" faction. Besides the mentalist understanding of memes, there is also a behaviorist understanding, which reduces memes to observable artifacts and behaviors as the focus of empirical work. In short, looking at all these different points of view, we can see there is currently no "standard codification of the concept" (Aunger, 2000: 7).

In addition, besides definitional issues, memetics faces serious empirical problems. For example, memeticists have yet to produce any empirical results, and on top of this, they need to design ingenious experiments to do so. Thus, memes and the means for finding them remain evasive. In a word, we must find memes and mechanisms for their replication, or the approach will fail. A later section will attempt to rectify this problem, looking at memes from a linguistic perspective. Moreover, as mentioned above, some memeticists claim an independent replicator dynamic exists for cultural change that is separate from human intentional agency, but we lack evidence for this as well. Lastly, this search for putative memetic bits of culture becomes even more complex because culture is a comprehensive totality of elements and thus intrinsically hard to define and reduce. Hence, despite the claims of some, memetics will probably never become the all-encompassing theory of culture, morality, and the minds that produce thoughts, artifacts, and moral

values. Nevertheless, memetics could benefit from small empirical successes set from a more modest agenda.

7.2.1 *Basic Background: Memetic Diplomacy, Pragmatism, and Mechanisms*

Besides these issues, memetics also faces diplomatic problems. For example, some meme research has alienated scholars in the humanities for failing to fully enjoy the vast literature on culture and mind (Blackmore, 1999). If this lack of careful scholarship were not enough, memeticists have made exaggerated scientific claims for their approach without empirical backing (Brodie, 2004). Unfortunately, this posturing and incautious reductionism will only impede our understanding of complex cultures, moralities, and the minds that make them. Hence, we can partly justify the allegation that memeticists have attempted to colonize the social sciences (Dover, 2001), as in some forms of social Darwinism, sociobiology, and evolutionary psychology. Thus, for memetics to succeed, scholars must do their homework, avoid posturing, and be careful when applying reductionism to culture and mind.

If these issues were not enough, some aspects of radical memetics confuse metaphysical presuppositions with empirical conclusions. For example, as we will see in more detail below, empirical science does not convincingly support the eliminative materialism of (Blackmore, 1999) and (Dennett, 1995) who, on this point, appear to conflate their science and philosophy. (Eliminative materialism sees common sense psychology as radically wrong, denying mental states such as beliefs). Moreover, this extreme approach can serve as an *unfounded* assault on any theory of morality that suggests genuine altruism can exist, because radical memeticists apply selfish "meme" theory to culture and morality. That is, altruistic behavior simply results from the influence of the selfish memetic replicators. However, as we have seen, a good case for real altruism still exists, and memetics may even help support it, but for scholars to take memetics seriously, its promoters need to avoid unfounded and tendentious claims.

On the brighter side, philosopher of science David Hull (2000) suggests that critics have set the standards too high for memetics -- higher than even now established

sciences could attain in their early stages. Thus, Hull urges memeticists to postpone definitional concerns and concentrate on getting empirical data because results will help resolve conceptual conundrums. Hull also compares memetics with linguistics, suggesting they could be mutually supportive, and he thus positively advises memeticists to develop their theory while testing it, suggesting specific goals. Hull suggests memeticists generate a theory of conceptual change; attempt to reconstruct conceptual phylogenies in memetic terms; and identify the mechanisms of memetic transmission, especially for replication and environmental interactions. However, Hull also cautions that meme researchers must remember that information theorists cannot distinguish between the replication of information and its implementation. This underscores the empirical side of the problem that memeticists have distinguishing between memes and memetic phenotypes, as Midgley has indicated (2001).

In contrast to these positive factors, the approach's most serious problems come from the "radical memetics" of Susan Blackmore (1999). To her credit, she has strongly influenced the popularity and scholarly debate on memetics. However, many consider her views unfounded. For instance, Blackmore claims that imitation is the only mechanism for memetic replication. However, there are problems with this. First, imitation in memetics is poorly defined as "copying a novel behavior or skill from another animal" (Blackmore, 2000: 56). Hence, if imitation *simply* means learning from seeing, then it transmits nothing of cultural import, and if imitation *broadly* constitutes every manner of communication, then it is too vague and thus meaningless (Plotkin, 2000). Second, imitation cannot account for the complex processes of how humans abstractly transmit concepts through recombination, invention, deduction, induction, and abduction. Third, other species imitate each other, and if so memes are not unique to human evolution as Blackmore claims. Fourth, counter-examples abound. The opposite of imitation may cause cultural trends -- where people behave a certain way because they *do not want to imitate others* (Conte, 2000). Moreover, Sperber (2000) claims localized instructions; information already in the system, instead of blueprints can transmit culture. Regarding language, Sperber says:

The similarities between the grammar and lexicons internalized by different members of the same linguistic community owe little to copying and a lot to pre-existing linguistic, communicative, and conceptual evolved dispositions. (Sperber, 2000: 172)

Lastly, Boyd and Richerson (2000) suggest Darwinian population processes -- not based on natural selection and replication -- may cause cultural evolution. Because of the rigor of Boyd and Richerson's work, memeticists cannot claim memetics *the only valid* Darwinian approach to culture, and so memeticists should accept a variety of approaches for the study of culture and mind.

7.2.2 *Background of Radical Memetics*

Perhaps "imitation memetics" is flawed only because it naively assumes imitation suffices for memetic replication when other possibilities clearly exist. However, in addition to this, Blackmore expresses problematic views. For example, she claims that without memetics, evolutionary theory would fail to account for the complexity of culture and mind. This is a bold claim because other scholars have applied non-memetic evolutionary principles to cultural change (Boyd and Richerson, 1988; Durham, 1991). Moreover, is it really a problem that natural selection might fail to explain culture and mind? As already mentioned, we may need other explanatory tools besides natural selection to explain language, morality, and culture.

Another bold assertion is the claim that memes are selfish replicators in that they "serve their own selfish ends, replicating wherever they can" (Blackmore, 2000: 54). This is the idea that memes replicate independently of human intent. Blackmore supports this logically by showing that some memes have a "viral structure," which includes a "copy me" instruction along with "threats" and "promises" that insure the meme's survival. However, she then unjustifiably reasons that memes imitate selfish gene selection, (Dawkins, 1976) where humans become meme machines, vehicles for the propagation of selfish memes. Blackmore boldly sums up her perspective:

We are neither the slaves of our genes nor rational free agents creating culture, art, science, and technology for our own happiness. Instead we are part of a

vast evolutionary process in which memes are the evolving replicators and we are the meme machine (Blackmore, 2000: 54).

This is why Blackmore thinks memetic selection, driven by imitation is what is missing from modern theories of human evolution. Thus, only memetics can enable us to explain our unique mental attributes, the complexities of our cultures, and the basis for our moralities. In short, Blackmore says humans exhibit a discontinuity from other species because we have been designed not by one replicator, but by a dynamic duo of replicators: genes *and* memes. That is, we are generated by genes and made by memes through a dualistic form of evolutionary design. However, if we emphasize discontinuity and duality, we have a problem because selfish gene selection (Dawkins, 1976), which forms the basis for the selfish meme concept, is essentially reductionist. But if we introduce the meme replicator, we expand entities, not reduce them. Moreover, memes are abstract, ideational, and non-physical entities (at least according to some memeticists), which replicate themselves in a world that, according to Blackmore, is ultimately and utterly materialistic. Hence, if memeticists are not careful, they will introduce a self-contradictory dualism into a reductionist understanding of culture and mind (Post et al., 2002).

7.3 Definitional Problems with Radical Memetics

Besides these issues, Blackmore poorly defines memes -- indicative of the fact the concept lacks clarity. First of all, she says: "Memes are best thought about not by analogy with genes but as new replicators, with their own ways of surviving and getting copied" (Blackmore, 2000: 61). This statement reveals a basic confusion: *are memes real or just metaphorical?* Simply put: since we have not found any memes yet, we cannot call them real. Since we have no evidence for memes, they are still at best just a metaphor on genes. This problem reiterates the lack of consensus about memes: some memeticists refer to the "meme analogy" and others call memes actual replicators. However, if memetics is only a new metaphor for talking about language, morality, culture, and the mind, then it does not give us novel insight into these subjects, but only a new way of talking about them.

Moreover, even more serious problems remain. First, scholars know how genes get copied, and they have mapped the human genome. However, memeticists give contradictory proposals for how memes might get copied, and this shows that we still cannot measure the transmission and reproductive fitness of memes.

Moreover, we do not know what constitutes memetic adaptation. If memes have a "viral structure," they may be more *structurally* fit than memes without it.

However, according to the theory, adaptive memes evolve alongside non-adaptive memes. For example, Blackmore discusses (2000) how adaptive fire building and hunting skills may evolve along with less or non-adaptive traits, such as body decoration and rain dancing. Nevertheless, we do not know if body decor and rain dances come from non-adaptive memes, and we can reasonably suggest that body decor enhances sexual attractiveness and thus reproductive fitness. Rain dances may allow the best dancers to show their reproductive prowess or increase their status. These traits might increase reproductive fitness thus making their meme adaptive. In short, given the nature of some evolutionary explanations (the just-so story type), we cannot easily conclude what constitutes memetic adaptation, and serious questions remain about what constitutes memetic replication, adaptation, and fitness.

Finally, we must focus on the idea that memes like genes "compete to get copied *for their own sake*" (Blackmore, 2000: 53). This claim asserts that memetic evolution occurs not in the interest of the genes, individuals, or groups, but that it proceeds according to the "interest of the memes" themselves (Blackmore, 1999). In memetic terms, "every human is a machine for making more memes, a vehicle for propagation and a resource to compete for" (Blackmore, 2000: 54). In shorthand, therefore, *memes want to use humans for their own reproductive success, and nothing else*. Blackmore discusses this form of shorthand:

We can say that memes are "selfish," that they "don't care," that they "want" to propagate themselves and so on when all we mean is that successful memes are the ones that get copied and spread, while unsuccessful ones do not. This is the sense in which memes "want" to get copied, "want" you to pass them on and "don't care" what that means to you or your genes. (Blackmore, 1999: 7)

However, though this is interesting rhetoric, we can see human intent in the way we constantly use language to describe our thoughts, values, and intentions. That is, we still have reason to believe that we are in the driver's seat of the meme vehicles. To be fair, when Blackmore says, "memes want to get replicated in humans," she means that the fittest memes get copied and spread and unfit memes do not. However, as Endler (1986), points out, this kind of thinking can be tautological. "If fitness is used in a careless fashion, it can lead to tautology, since by definition the most fit survive or otherwise do better than the less fit" (Endler, 1986: 33). Moreover, the following echoes this same concern about memetics:

The ability to define fitness independently of what evolves saves the concept of natural selection from being a tautology. For the meme concept to escape the same problem, we must define cultural fitness independently of what evolves. If the first four notes of Beethoven's fifth is a powerful meme only because it is common, we have achieved no insight (Wilson, 1999: 206)

Hence, we can see the tautology problem in the following line of circular thinking where memetic fitness is not defined independently of what evolves.

1. The fittest memes survived and reproduced.
2. Why did these memes survive and reproduce compared to other memes?
3. They survived and reproduced because they possessed more adaptive and profitable memetic traits than memes that didn't survive and reproduce.
4. How do we know that these traits were the most profitable and adaptive?
5. We know these traits were most profitable and adaptive because the memes that possessed them survived and reproduced.

Memetic theory does not have to be trapped in tautology. We must define fitness separately from what evolves. If we find the specific fitness enhancing characteristics that caused memes to survive and the fitness diminishing traits that decreased the survival potential of other memes, then fitness is separate from survival.

Besides this, memeticists need to be careful with the metaphors such as "in the interest of the memes; what the selfish memes want; and memes want get copied." This shorthand reveals a problem. Memetics not only prescind from teleology (avoids questions of purpose), as evolutionary theory generally does, but it also claims that minds and cultures exist as fully dysteleological entities produced by non-teleological processes. Nevertheless, memeticists often speak in teleological

shorthand, and though this shorthand is not inherently problematic, I argue below that this tendency reveals another set of problems.

7.4 Misguided and Misinformed Aspects of Memetics

We can summarize radical memetics as follows. Memetics is a Darwinian dysteleological theory of culture and mind that often employs teleological language to describe putatively non-teleological processes and events. If we personify nature or speak of it teleologically, this is not wrong in certain contexts, even though literary scholars call this kind of anthropomorphism the pathetic fallacy (attributing human emotion or responses to nature). However, if our methodology dictates we explain cultural and mental processes in dysteleological terms, this still does not mean culture and mind are non-teleological. This could be a fallacy of composition, inaccurately grouping our method and subject in the same non-teleological category. We could view culture as designed by and for the memes, and view humans "not as rational free agents" whose minds do not and cannot create culture, if this were true to experience. However, in simple terms, many factors confute (or at least challenge) the "dysteleological theory" of mind and culture, including human experience, perception, common sense, and the empirical pointer of language usage.

Nevertheless, intelligent people espouse this idea, which raises the question: "why would anyone hold to a theory that is so drastically counter-intuitive and that strongly clashes with human experience?" First of all, science has had tremendous success prescinding (detaching itself) from the question of telos in nature. Because this methodological naturalism has proven so successful, it seems logical to continue to use it in our study of culture and mind. However, we need to distinguish methodological naturalism (MN) (Daniel, 2005) from teleomentalism (TM) (Colin, 2004). MN is a methodological thesis about how we study nature, and we can distinguish MN from physicalism, which is a metaphysical thesis about nature. TM in its strong form holds that teleological claims in biology are mere metaphor and thus we can eliminate any references to teleology in biological science. We may see strong traces of physicalism in TM in that TM is a

physicalistic view that is also more metaphysical in character than it is methodological.

With MN we will look at cultural and moral code transmission without considering issues of purpose. That is, we will try to find the aspects of this transmission that occur separately from human purpose and intent. However, TM in memetics sees purpose as trivial or unreal, attributing all power to memes. That is, MN just ignores purpose to see what can happen without it, but TM makes MN a metaphysical principle about what is real and can be known. Some scholars seem to conflate these so that MN takes on a strong epistemological status and becomes the only way of knowing. Ideally, however, we should use a method with two types of lenses -- one that lets us see what happens without purpose, and another that lets us see the potential effects of purpose and human intent in the process of cultural transmission.

Thus, we should remember that when memeticists say, "we are non-free, non-rational agents," they are not making an observational claim emerging from naturalistic methods. Perhaps they are stating a theory that we might test, or they are letting pre-empirical beliefs influence their scientific claims, or at worst they are simply making metaphysical statements and unjustifiably placing them under the authority of science. In short then, memetics should benefit from a methodology that filters out telos in both its explicans and explicanda. However, if memeticists claim factually that mind and culture are non-teleological, then we might want to check their presuppositions for overdose of eliminative materialism (or TM), which claims that notions of belief, experience, and sensation are fundamentally mistaken. A more moderate and balanced approach will say that we are not looking for purpose, but how culture might get transmitted separately from purpose. Better yet, we might look for ways that culture and morality get transmitted with and without purpose.

Additionally, non-teleological memetics contradicts basic human experience, perception, and common sense. This comes to the fore when we consider Blackmore's memetic concept of self, or "selfplex." Blackmore boldly calls this self-concept a terrifying fact, and she describes it polemically and nihilistically:

In fact we know that selves are a myth. Look inside the brain and you find only neurons. You do not find the place where 'my' conscious decisions are made. You do not find the thing that lovingly holds all those beliefs and opinions. Most of us still persist in thinking about ourselves that way. But the truth is -- there is no one in there!" (Blackmore, 1999: 43-47)

Blackmore overstates her polemic by saying we *know* that selves are a myth, as if memetic research confirms eliminative materialism or physicalistic teleomentalism. Perhaps she confuses her presupposition for her conclusion. More importantly, other valid theories of the self exist that do not simply nullify it. For example, Polkinghorne (2003: 22) suggests the self forms the "information bearing pattern of the body." See also (Richards et al., 2002) for an interesting debate on the subject from various perspectives. Perhaps Polkinghorne's idea fits better with experience than reductive theory, for seeing "only neurons" fails to account for this coherent and continuing body of information inherent to human experience.

We may not be able to find a person by poking at their gray matter, but this one-dimensional view hides a more important question of how personhood can emerge from the psychophysical processes in the brain. If we only acknowledge the gray matter in the brain, this is like only acknowledging the ink and paper in this chapter. We thus fail to see the semantic meaning emerging from the words and sentences -- and the meaningful patterns of personhood that emerge from the brain. Perhaps Blackmore would suggest that only the memes are there, but we would be more consistent to propose that this information pattern constitutes a self. Moreover, excising the self from mind, culture, and morality is no more parsimonious than including the self, because we simply replace the self with memes.

Moreover, we can easily show how "memetic selflessness" clashes with common sense. Though common sense could be wrong, Blackmore must seriously engage the issues raised by common sense and experience instead of just preemptively dismissing them. I offer the following scenario (If Memes Had a Mind of their Own) to illustrate the problems with 'selfless' memetics.

7.4.1 Scenario: *If Memes Had a Mind of their Own*

In the selfplex view -- no self, only memes -- who care only for their replication influenced the writing of this thesis. Moreover, if another meme-controlled non-person were to read this, his non-self would just be a physical entity submitting its-non-self to memetic processes controlled by and for memes. Therefore, a non-person wrote this thesis, and non-persons will read it in a purely self-less and non-personal process. Moreover, if this thesis were read at a conference, no particular person would hear it, and no-person would respond to it personally because persons and personalities do not exist, only memes do. Hence, if some non-person were to disagree with the contents of this paragraph, it would only be the *appearance of a person* who disagrees because non-personalities cannot agree or disagree in any personal sense. Actually it would only be memes in competition and 'disagreement' with other memes, contending for survival in human neuron space.

Moreover, if a non-person-reader were to become enraged over the contents of this paper and kill the non-person-writer; the fact would emerge that no person killed no person. That is, if memes in the reader-entity could physically arouse rage in the reader-entity, animating "it" to kill the writer-entity, then only memes would be responsible for this murder as they competed for neuron-space to ensure their reproductive fitness. Of course, in most societies, law requires that such non-persons control their non-selves, or their memes, even though, according to radical memetics, *no such selves* exist. Moreover, since the writer-entity is now dead, we can conclude this thought experiment by saying first, the bad news (or is it good?): *no person will mourn this non-person's death* and the good news: *some memes may have increased their chances for survival through this homicide*.

However, there is a complication. If the memes that dominated the writer-entity found out about the expiration of their neuron-space, they might arouse a flow of justice memes in the relevant human population. These memes might cause non-persons to attack the meme vehicles that caused the death of the writer-entity. Hence, memetic justice could be attained, stopping or decreasing the killing of human entities that contained this writer-entity's memes, and this would explain in memetic terms why murder is prohibited in almost all cultures: *it maddens the memes*.

7.5 Beyond Memetic Madness

Though this memetic allegory seems outlandish, and it does not represent all forms of memetics, eliminative materialism logically leads us in this way. Sommers and Rosenberg (2003) claim that Dennett (and therefore Blackmore) do not work their ideas to their logical end. To be consistent, we should say that ethical beliefs "where they exist, they are false" (Sommers and Rosenberg, 2003: 655), and thus

morality emerges from mindless and purposeless forces, meaning that "There is nothing left of morality to explain" (667). However, this nihilism, at its logical dead-end, explodes the very path it creates, undermining epistemological warrant for any beliefs, including ones that support memetics and moral nihilism. That is, if eliminative materialism erodes belief in morality, it also erodes belief in itself and everything else. We end up with no reason to believe nihilism and the memes that encode it since our presuppositions force all our ideas into nullity. This kind of memetics self-destructs.

Therefore, moral responsibility, free will, and culpability become problematic in this context, and such moral concerns are central to the focus of this thesis. Though we might want to remain open-minded about this kind of memetics, it appears to be seriously trapped in fatal logical flaws and unworkable fatalism (Midgley, 2001). Moreover, this kind of memetics even contradicts its parent theory of genetic selfishness. That is, some memes could promote their own survival by getting their human vehicles to sacrifice their own fitness for the sake of the memes. Though Dawkins mentions this kind of insurgence against selfish genetic replicators, he does not at that point take the meme's eye view.

Nevertheless, memes could sacrifice their human vehicles to keep their memes alive in others. For example, democratic memes see their future fitness in danger from an attack of fascist memes. Thus, democratic memes mobilize human vehicles to attack the human vehicles of fascist memes, and in the process many "non-persons" bearing the democratic meme die in order to end the dangerous advance of fascist memes. Here pure altruism, the sacrifice of individual genetic fitness, emerges for the sake of the memes. This scenario might actually explain altruism, where humans use memes (or memes use humans) to motivate real altruism. However, the problem not only comes when we remove human personhood from the picture, but also how democratic memes could independently replicate themselves -- separate from human purposes. Maybe memes can possess us, but if we claim we are totally meme-possessed, this may result in many logical flaws and factual mistakes.

Dawkins (1976) does not seem to hold to this eliminative materialism, which eradicates personhood. He claims *we* can and should rebel against the tyranny of the selfish replicators; that is, memes do enable *us* to rebel against genes, but Blackmore (1999: 246) counters by saying that there is "no one to rebel." Perhaps Blackmore is serious, or perhaps this is some kind of academic posturing. However, because empirical evidence does not support the conclusion that personhood does not exist, and because this claim fails to match human experience, and because it implies extremely counter-intuitive consequences and ethical dilemmas, the argument must be flawed. Moreover, as previously mentioned, it is also clear that Blackmore is overstating her case, confusing her metaphysics with science, for memetics and the unproven existence of memes cannot prove eliminative, dysteleological philosophy. Hence, as a psychologist, Blackmore unfortunately uses philosophy, not data to drive her conclusions. Nonetheless, at least we can commend her for correctly reasoning from her presuppositions, though it seems to lead to a kind of scientific madness eerily reminiscent of the "Un-man" in C. S. Lewis' science fiction novels (2003).

7.6 The Potential of Memetics in Metaphysics and Science

Up to this point, we have seen many problems with memetics, and we have not had space to see the way Dawkins and Blackmore ineptly discuss religious memes, (see (Wilson, 2002; McGrath, 2004)), which are very important to the subject of morality. Because of all these problems, many scholars have written off memetics (Bloch, 2000; Kuper, 2000; Sperber, 2000; Dover, 2001; Midgley, 2001); thus, perhaps we should reject the whole project, call it a lost cause, and move on. However, if we carefully note the problems and accept positive support of some leading thinkers (Durham, 1991; Dennett, 1995; Hull, 2000; Aunger, 2002; Balkin, 2003), perhaps we can test a modest memetic approach. Hence, current problems may help us redefine it in the right direction. For example, we can focus on (1) ways to find an independent replicator dynamic for memes (the idea of memetic fitness), (2) exploring how memes might relate to evolution and morality.

7.6.1 Neuromemetics

As a linguist, I find it interesting that apparently no one has done what Hull (2000: 42) recommends: conceive of memetics as a version of linguistics based on "selection as it functions in evolutionary biology. This idea of "lingua-memetics" relates to language and morality in four ways. If memes exist, they could (1) form a vital element of protolanguage, (2) contribute to the evolution of human groups in protolanguage or fully complex language, (3) influence the evolution of altruism, and (4), possibly facilitate group selection or more radical forms of altruism.

First however, for any version of memetics to become viable, we must amend it by correctly specifying and finding an authentic mechanism for memetic replication, by rigorously defining and discovering real memes, and by finding evidence for an independent replicator dynamic. To date, Aunger makes the best effort to solve these problems, but as I show below, he still fails. For example, Aunger (Aunger, 2002: 73) stipulates that true memetic replication will also fulfill four "replication criteria" of causation, similarity, inheritance, and duplication.

<i>Causation:</i>	The source must be causally involved in the production of the copy.
<i>Similarity:</i>	The copy must be like its source in relevant respects.
<i>Information transfer:</i>	The process that generates the copy must obtain the information that makes the copy from the same source.
<i>Duplication:</i>	During the process, one entity must give rise to two (or more).

Aunger also makes numerous important claims for his neuromemetics, which both support and contradict the idea of lingua-memetics. Aunger is to be commended, for he has helped provide the best critique of memetics to-date in the edited volume *Darwinizing Culture*, (Hull, 2000). Though authors suggested positive ideas for salvaging memetics, many of them decried the grave problems and seemed to announce its doom. However, *The Electric Meme* (Aunger, 2002) answers many of the criticisms and provides the most serious attempt at memetics to-date.

To clarify how memetics might enlighten our understanding of language, evolution, and morality, we need to focus Aunger's views (2002). Aunger claims that "Only brain states have the necessary qualities to replicate: They can cause

similar entities to arise through information transfer" (2002: 324). Thus, artifacts and behaviors fail the replicator test. Artifacts fail as replicators because Artifact A plays little or no causal role in the making of Artifact B. That is, a newly completed car does not make the next car on the assembly line, and there is no information transfer between cars on the assembly line. Specialized kinds of artifacts such as computer viruses may qualify as replicators, but only in a limited sense because they cannot explain the totality of culture, or for that matter the relationship between morality and language.

Regarding behaviors, Aunger demonstrates they cannot qualify as replicators, because behaviors, like signals, are not duplicated and because tracing the path of a behavior "does not identify distinct evolutionary lineages" (2002: 173). Thus, Aunger concludes that if we are going to find a meme in the brain, it must be in "the *state* of a node in the neuronal network" (2002: 324). He concludes, "the important point is that a state, unlike the physical network itself, can be duplicated elsewhere in the network" (325). Moreover, these meaningful patterns of electro-chemical matter can be replicated extremely rapidly across the neuronal network. Additionally, they are independent of genes because though they reside on a physical substrate produced by genes, they are separate from it because they can be produced by other stimuli, such as other memetic products.

In this way, Aunger attempts to physicalize memetics in the brain. Nevertheless, this does not begin to explain how memetics might begin to help us understand culture, and the relation between language, evolution, and morality. However, the following clarifies the broader aspects of his approach:

... in neuromemetics, which covers the first phase of meme evolution, information transfer is satisfied by the production of a signal spanning the gap between one neuron and the next; replication is based on the mediation of signals *within* organisms. Social memetics (the second phase) depends on signal-mediated replication between organisms. In coevolutionary memetics (the third phase), information transfer becomes even more indirect. First, a meme-inspired behavior produces, or modifies an artifact, which in turn does the job of producing or modifying a signal that eventually reaches a new host. It is a case of *artifact- and signal-mediated* replication between organisms. (Aunger, 2002: 329, emphasis in the original)

7.6.2 *The Flaws of Neuromemetics*

Hence, memes are electro-chemical patterns residing on a node or multiple nodes in the brain, and they replicate between brains first through signals and then through artifact-signal mediated replication. This is interesting, but it raises the question that if memes are simply node-states in the brain, what is the difference between a meme and any thought or memory? If words encode on a node-state in the brain, does this mean that every word we remember constitutes a meme? Some words may be ideal examples of memes, but brains encode all information in brain node-states, such as memories of smells, sights, sounds, and sensations. However, it seems unlikely that my visual memory of a cloud formation, which I saw today, and which is encoded in a brain node-state in my brain, constitutes a memetic bit of culture that will be replicated dynamically and independently of my intentions.

Conversely, even if we claim that a memory of a cloud formation is a meme, (it could become a memetic artifact in a painting), what about my memories of these things: the smell of fire, the sound of blowing wind, or the feeling of cold water? All memories must encode in brain node-states, but do they all constitute memes? Neuromemetics may have trouble distinguishing between sense memories like these and memes, and *therefore a meme must be more than just the state of the node on a neuronal network*. Memetics must distinguish between memes and all kinds of thoughts, mental concepts, or sense memories. If memes are just any kind of mental concept, then memetics is still vague and unhelpful. Perhaps there are two kinds of brain node-states: non-memetic and memetic, and the memetic kind experiences replication. If we can differentiate between sense memories and memes that attain memetic fitness, (a catalytic state) that replicate in degrees independently of human intent, and not just within the brain but also between organisms, then memetics will attain a fascinating explanatory status.

If we can locate memes in brain node-states, this will show that memes could replicate within brains independently of human intention. Nevertheless, intentional involvement (study, for example) could increase the effectiveness of memes replicating to different nodes within brains. After all, we can enhance our memory of "memes" through efficient and effective effort. However, in order to distinguish

between memes and every word, sight, smell, sound, or sensation, we must specify the memetic condition of the node-state in the brain.

If memetic-node-states actually exist, they may still replicate within brains in the same way as non-memetic-node-states. However, if non-memetic and memetic node-states both exist, then how can we qualitatively or quantitatively distinguish between them? Unfortunately, the difference here appears qualitative rather than quantitative. Moreover, though an independent replicator dynamic might act at the neurological level, the social level is more significant. The striking memetic claim concerns whether culture can replicate independently of human intent. That is, if someone hears a memetic message, will memes replicate in his brain independently of his intent, and will memes replicate through him to others in the same way? Do senders of memetic signals between brains communicate intentionally or by the independent or catalytic action of memes in their brains, or do memes replicate in a combination of catalytic and calculated action?

In neuromemetics, Aunger shifts the "perspective from social to inter-cellular communication" (2002: 331), and this causes us to look for physical memes where they will most likely experience catalytic or independent replication. However, since memetics aims to explain culture, even if we find memetic-node-states in the brain and can distinguish them from non-memetic-node-states, we must still find memetic lineages and discern how they go between brains and artifacts. Aunger proposes (1) social memetics: signal-arbitrated replication between organisms, and (2) coevolutionary memetics: signal and artifact-arbitrated replication between organisms. Though these proposals may explain how memetic evolution might work, these two levels of memetic processes appear much more central to memetic theory than Aunger suggests.

It's true that the life cycle of a meme sometimes involves the use of signals that work at the inter-organism scale, but memes *per se* are not making the journey between hosts. (Aunger, 2002: 331)

If memetics truly concerns culture and if memes actually exist, then they do not just *sometimes* use signals at the inter-organism scale. The cultural process centers on the inter-organism plane: it is social. This means that the central focus of

memetics must deal with how and why memes travel from brain to brain. Memetics began as a theory of cultural evolution, not as a theory of consciousness. If we focus on brain states, this may turn memetics into a physical science; however, this may be overly ambitious and unrealistic. Moreover, we can make a strong case to focus mainly on the social level of signal-mediated memetic replication because understanding the memetic journey between brains is central for us to explain culture memetically.

Aunger's memetics also faces another problem because he claims that it is Darwinian and not Lamarckian, in contrast to Gould's claim (2001) that cultural change is Lamarckian. Aunger asserts that memes do not make the trip between hosts, but signals as instigators and not phenotypes make the journey and then instigate the memetic state in another host. This means, "If signals are instigators, then replication also doesn't need to involve construction, just *conversion*" (Aunger, 2002: 242, emphasis in the original). Thus, the signal flips a neuron in the target brain from one state to a new state that exists in the original host. Aunger adds:

However, producing an infectious state in the receiving brain is not enough: it has to be the same state as produced the signal in the first place to constitute the next link in the memetic lineage. (Aunger, 2002: 242)

Aunger's theory of social memetic replication is interesting, but we will find it very difficult to demonstrate that a signal can instigate the same neuronal state from one brain into another brain. That is, if I say "banana," and you hear and repeat it, how can we know that this word will form the same neuronal state in both brains? Neuronal encoding could be different for the same word in separate brains. In any case, it will be very hard to figure out exactly what the same node state actually is; thus, this may turn out to be another untenable concept for memetics.

Moreover, memetic transmission may transcend Darwinian processes, and neuromemetics does not predict this. In fact, because of pre-existing states in a recipient's brain, signals may intermingle with these states in ways that instantaneously convert nodes into different states than the source meme in the original brain. If this occurs, the *process may create new memes rather than*

replicate them. This problem is similar to imitation in the lesser and flawed versions of memetics. As imitation cannot account for some abstract forms of transmission, signal instigated replication cannot explain what happens to memes when they encounter the combinatory, inductive, or abductive abilities of the target host.

Thus, memes can derive into different memes during the process of signal-mediated social replication. Moreover, Conte's (2000) argument against imitation memetics applies here too. People may resist memes; willingly or unwillingly conflate memes with other memes; they may confute memes, or they may misconstrue memes. They may also oppose particular memes and reinterpret them in ways that fit within their pre-existing system of thought. Hence, though some signal-instigated social memetic processes may result in the simple conversion of the target host's neuronal node state, there is no guarantee that signal instigation will always result in this straightforward memetic replication process.

7.6.3 *Linguistic Problems with Neuromemetics*

Aunger's memetics also faces problems regarding how he looks at language. He suggests that a memetic spike in the initiating host stimulates a motor neuron that produces a social signaling behavior, such as speech, that a target organism consumes. These social signals then replicate the same neuronal node state in the target organism. In this view of communication, Aunger avoids the inferential approach as seen in Grice's (1975) maxims that require interlocutors to infer intended meanings cooperatively and accurately because signals do not contain everything a sender wants to communicate. Aunger thus concludes that inferential dialog *cannot* explain signal mediated memetic communication.

The consequence of adopting an inferential approach, however, is to separate the study of communication from the rest of science and make it something of an art form: the subjective study of human intentions (Aunger, 2002: 261)

However, if we remove inference from our understanding of human communication, this may render it meaningless. Disregarding inference might show us something novel about communication, but if inference plays a significant role in communication, then disregarding it will mislead us. Moreover, Aunger

claims the inferential view assumes cooperation, but this is not quite accurate, for we can easily turn Grice's cooperative principle on its head and make it the *uncooperative principle*. That is, interlocutors can communicate non-cooperatively, inferring meaning incorrectly about each other's intended meanings.

Though we might think that non-cooperative communication will die out, a more subtle result obtains. Non-cooperative communicators simply go on misunderstanding each other; thus, their communication may thrive in antagonistic argument and debate. For example, two leaders in a controversial public debate may continue arguing with each other for years, misunderstanding and misconstruing each other's positions, and they may misunderstand each other because they distrust each other. Thus, if Debater A says, "we want to help the little guy," Debater B thinks, "How condescending! He wants to control and use the public." However, Debater A may actually want to help people, and may have no desire to manipulate them, so an inference based on distrust skewed Debater B's understanding. These Debaters share a mutual interest in an issue and show a kind of pseudo-cooperation, but their distrust of each other causes them to make wrong inferences, and these inferences strongly influence their communication. In short, the uncooperative principle will promote miscommunication by skewing the inferences of the interlocutors.

The cooperative principle says people will be "informative, truthful, relevant, clear, unambiguous, brief, and orderly" (Pinker, 1994: 228). However, when they do not trust each other, interlocutors may expect each other to be uninformative, untruthful, irrelevant, unclear, ambiguous, verbose, or disorderly. Genetically selfish conspecifics might presuppose non-cooperation and infer meanings in this light. However, this thesis demonstrates that not all behavior is genetically selfish; real altruism exists, and communicators can infer meanings in both cooperative and non-cooperative ways. We can communicate altruistically or selfishly. In sum, if memetics can help us understand culture and moral transmission, it cannot neglect the key role that *cooperative* and *non-cooperative* inference plays in communication.

In addition, Aunger may under-estimate the altruism of communicators and over-emphasize their selfishness. That is, though conspecifics may host a "*parasitic replicator with its own evolutionary interests*" (Aunger, 2002: 265, italics in the original), these conspecifics may not always have different genetic and social interests in contrast to what Aunger states. "Indeed overlapping interests should be considered an unusual case since genetic endowments and socioenvironmental situations are rarely the same for different organisms" (2002: 261). There will be selfishness and divergent interests, but language provides a mechanism that helps humans cooperate. Moreover, memes that promote altruism and group cooperation may reap fitness benefits, if these memes help a group survive.

In fact, memes could exist that promote cooperation and morality. Dawkins (1976) claimed that memes would help selfish organisms rebel against selfish replicators. That is, with language, (or perhaps with lingua-memes) we can promote cooperation, and we can also infer meanings cooperatively. Our study of communication may become fuzzier and messier than we want, but the truth about communication may actually be a fuzzy and messy truth. To the point, selfish gene theory does not refute inferential cooperation in communication. If we disregard cooperative inference (as Aunger does), we overstress the principle of parsimony and fail to realize the complexity of culture, communication, and morality.

To compensate for the problems with inferential communication, Aunger refers to the Chomskyan view about the poverty of the stimulus. That is, an innate grammatical instinct may prewire us to learn language despite a paucity of input -- to learn more language than the amount of linguistic stimulus would predict, and this instinct may also help us reconstruct meanings as well. Aunger says, "to correctly interpret the sender's intention, the receiver must fill in the gaps in the message and thus reconstruct what the sender intended to say" (Aunger, 2002: 247). Aunger claims that innate structures in the brain, as proposed by evolutionary psychology, would "provide everyone with the same mental code-decode mechanism" (2002: 248). This would regularize communication and allow interlocutors to reconstruct meanings correctly. If we can reconstruct meanings correctly, memes will replicate between brains because evolutionary and biological

constraints would foreclose the kinds of inferences that minds can make about memes. In the end however, it is not clear how this differs from inferential communication. Moreover, though mental structures may help determine cultural and linguistic universals, this does not exclude intentional, deductive, inductive, abductive, and other forms of inferential abstract reasoning. Thus, instead of simply replicating in brains, memes could actually spark new flashes of insight that transcend the memetic process.

7.7 Lingua-memetics: Fitness-enhancing information in a memetic state

Up to now, we have seen numerous problems with memetics -- summarized here:

- Memeticists have found no evidence for an independent replicator dynamic for memes, nor any mechanism for their replication.
- Memetics sees cultural change as Darwinian when it may be Lamarckian.
- Memeticists have neglected the existing literature on culture and mind.
- Memetics attempts to reduce culture to bits, but the complexity of culture may make this impossible.
- Memetics avoids (or denies) questions of purpose or intention in a world full of purposeful or intentional agents.

Though this appears bleak, we may cautiously hope for a limited version of memetics to give us insight into culture. Moreover, because language supports morality, *if memes exist in language*, lingua-memetics might help us understand altruism and morality. In a linguistic theory of memetics, the human language capacity would form the mechanism for memetic replication, not only as Aunger suggests by giving people the same "code-decode mechanism" that helps us fill in the blanks from context, but in other ways as well. UG also provides the biological specialization that designs our grammar acquiring capacity and the commonalities of human languages. Most importantly, symbolic languages provide the fundamental means to transmit culture and cultural values. Hence, though our prewired brains give us language ability, symbolic languages themselves must encode and transmit memes. That is, "humans learn not only by imitation and experience but also by accessing information from cultures that are encoded in symbolic languages" (Goodenough and Deacon, 2003: 809).

Symbolic languages encode information, but factors intrinsic to the meme itself, the way it is expressed phonetically and grammatically may increase its chances of catalytic replication. That is, linguistic traits may cause ideas to replicate more effectively, turning these ideas into memes, or causing them to attain a memetic state. For example, oft-repeated proverbs might attain a memetic state because they are pithy, succinct, and easily memorized; moreover, they might also encode semantic information that could enhance the fitness of their hosts and that requires the hosts to replicate these memes. Such memes would be memetically fit, not because they survive, but because they have particular set of traits that help them to survive and replicate catalytically and independently of human intention, a set of traits that other ideas do not have. Thus, we define memetic fitness separately from what memes evolve and survive because memetic fitness can come from linguistic traits encoded *in and onto* the meme that are also detachable from the meme. These traits are described in greater detail in the next sub-section.

Incidentally, lingua-memetics focuses on factors that cause ideas to coalesce into a memetic state and replicate through language, and this avoids the problems in Aunger's proposal regarding brain states and neuronal nodes. We will find it very difficult, if not impossible, to physically detect memetic brain states. For example, people process the same kind of information in different parts of their brains. Thus, meme X may be located in different places in different brains, so until researchers have the technology to scan brains for specific information, we cannot find physical memes in the brain. Moreover, if we avoid the serious problems of trying to detect memes in neuronal node states, we will also avoid conflating memes with other non-memetic brain states storing recollections of sights, smells, sounds, and sensations.

7.7.1 Looking for Memes in Language

With these issues in mind, the following section outlines 16 factors that may help us find actual lingua-memes -- concepts, which attain a memetic state and experience varying degrees of catalytic replication through language. Dawkins (1976) and Blackmore (1999) already have suggested variations of factors 3-7. To reiterate, these ideas *become memetic* because phonetic, grammatical, and lexical

factors make them *cognitively sticky* or more memorable, thus increasing their chances for catalytic replication. Moreover, these phonetic, grammatical, and lexical factors attach themselves to content factors that may increase both the fitness of the hosts and the memes as well. Thus, memes may possess *some or all* of the following traits.

— *Content Factors*

1. Memes may contain fitness enhancing information that teaches memetic hosts how to better survive, e.g., how to care for offspring, how to maintain health, or how to efficiently get food.
2. Memes promoting fitness may be stated as pithy proverbs that express cultural or moral universals (Strungaru and Schiefenhövel, 2002).
3. Memes may contain their own replication or "copy me" instructions to ensure their fecundity and longevity.
4. Memes may contain correction or "repair me" instructions to ensure fidelity of copying.
5. Memes may contain redundancy or "reiterate me" instructions to ensure constant reactivation of memetic states in their hosts.
6. Memes may contain "obey me" instructions, supported and enforced by verbal and social promises, threats, rewards, or punishments.
7. Memes may contain "believe me" instructions that would increase in catalytic potential if connected to metaphysical beliefs about deities, demons, ancestors, or heaven and hell.
8. Memes may contain "be altruistic" instructions, promoting altruism that enhances group fitness and ensures the survival of these memes in the group.
9. Memes may contain "be groupish" instructions to help identify and demarcate the in-group from the out-group, helping separate insiders from outsiders, believers from non-believers, and the initiated from the uninitiated.

— *Phonetic and Acoustic Factors*

10. Memes may coalesce with poetic factors, rhyme, assonance, consonance, alliteration, and meter that may increase memorability.
11. The memetic status of these concepts may also increase through musical factors: melody, harmony, and rhythm.

— *Grammatical and Lexical Factors*

12. Memes may find expression in concise (focused), succinct (compressed), and pithy (substantial) phrases that increase memorability and the chances of catalytic replication. "Do as you would be done to" may be more memorable than "Treat others in the same way that you want them to treat you."
13. Memes may be reducible to holistic phrases that are highly contracted, and where mention of just part of the phrase may elicit the full phrase in the minds of the interlocutors. The phrase "Do unto others..." may serve as shorthand for the whole Golden Rule.

— *Social and Functional Factors*

14. If people believe in a group's memes, they may garner that group's acceptance, but if they reject those memes, the group will reject them. This places pressure on individuals to believe the meme.

15. Memes may be false beliefs, but they will likely possess strong functional or adaptive truths that help them maintain their appeal in the face of strong counter-evidence (Wilson, 2002).

— *Summary*

16. Moral memes enveloped in the above conditions may attain an internalized, second nature neuronal encoding, like conscience. In this sense, memes could *possess their hosts* and could likely replicate through language in varying degrees dependent and independent of human intentions.

In short, numerous factors and pressures may coalesce to make information memetic. This list is not exhaustive, but it outlines a reasonable approach to what kind of information might attain a level of memetic fitness. Moreover, though we may need to more rigorously define each of the above traits, none of them seems beyond the reach of some kind of tangible or empirical testing. For example, we can prime subjects with only part of a proverb to see if they can recall the entire phrase, and we can see if subjects can recall proverbs that contain more of these traits than phrases that possess fewer of them. The most important question concerns whether we can demonstrate that a phrase can replicate dynamically between brains in varying degrees independent of human intent -- thus attaining memetic fitness.

Neuroscientists will face the potentially impossible task of distinguishing between memes and all other recollection states of smells, sounds, sights, and tastes that undoubtedly replicate between nodes *within brains*. However, if we want to explain culture in memetic terms, we must understand how memes replicate dynamically *between* brains, so understanding cultural transmission is the most important and realistic goal of memetics. Thus, we must describe the kinds of memes that potentially have the most power over us, and then we must look for ways to evaluate whether these memes can replicate independently of our intentions. In short, our current capabilities and realistic goals should lead us to look for fitness enhancing information in a memetic state. The above traits tell us what memetic information could look like, and thus we should focus specifically on how such information might replicate independently of human intention. If we can demonstrate that this kind of information replicates catalytically, we can confidently call it memetic.

7.8 Potential Examples of Linguistically Based Moral Memes

Therefore, how can researchers find ideas that may possess humans and that we replicate, at least to a degree, independently of our intention? We can search for a concept that matches the above traits. Since the Golden Rule (GR) has stood as a significant theme in this thesis, it is a logical choice for one of our putative moral memes. For trait 1, we can argue that the GR contains fitness enhancing information (it teaches us how to survive in a group). Moreover, though adherence to the GR might cause one to risk fitness through altruistic behavior, under certain conditions, the behavior would also provide adaptive benefits for group survival. For trait 2, the GR teaches fitness to individuals on how to survive in groups, and it may also promote group fitness by encouraging altruism in individuals. Moreover, the GR expresses a cultural universal that we may state in a pithy proverbial form.

Trait 2 concerns whether a particular meme in question exists as a cultural universal. As previously indicated, versions of the GR exist in Buddhism, Christianity, Confucianism, Islam, Hinduism, Judaism, Taoism, Sikhism, and Zoroastrianism, so it obviously stands as a cultural universal. For our focus here, we want to ask why this is the case. Is it possible that our biology predisposes us to easily learn the GR? For example, does our natural capacity for empathy influence the ease with which we can learn the GR? Does the way we linguistically render the GR help it attain a memetic state, and how could we empirically try to answer this question cross-linguistically? Again, there are more questions than answers, but if we can formulate these questions into tangible research projects, perhaps we can begin to see if some concepts in language attain memetic states.

For traits 3-9, we will look at the following Deuteronomic text and surrounding context, which does not explicitly state the GR. However, Deuteronomic laws state many GR type rules and provide good examples of ideas that may have attained a memetic state. First of all, for trait 3, this text contains "copy me" instructions.

These are the commands, laws, and rules the LORD your God commanded me to teach you. Obey them after you enter the land and take possession of it. As long as you live, you, your children, and your grandchildren must fear the

LORD your God. All of you must obey all his laws and commands that I'm giving you, and you will live a long time... ..Repeat them to your children. Talk about them when you're at home or away, when you lie down or get up. Write them down, and tie them around your wrist, and wear them as headbands as a reminder. Write them on the doorframes of your houses and on your gates. (Deuteronomy 6:1-2; 7-9, New International Version)

For trait 4, though this text does not contain "repair me" instructions, the Jews had a rigorous scribal tradition with strict rules for the fidelity of copying of their scriptures. For trait 5, these injunctions above clearly contain "reiterate me" instructions -- so that the meme will be *repeatedly replicated*. Moreover, though the GR does not contain a "reiterate me" instruction, the context for the Judaic moral text with its GR type rules issues about 20 commands that tell adherents to reiterate these instructions carefully and constantly to others. For trait 6, the above text has "obey me" instructions that were backed up by promises, threats, rewards, or punishments. For example, stealing violates the GR, and if you steal your neighbor's cattle, you will be ostracized (lose group protection), or you will be made to pay recompense (lose resources). For trait 7, this text comes from a religious tradition that increases the weight of these memes by associating them to beliefs about a deity who backs up these instructions with supernatural power. Hence, the cultural concepts that surround the GR provide "copy, repair, reiterate, obey, and believe me" instructions, which may help catalyze the GR (or GR type rules) into a memetic state.

Regarding trait 8, Chapter 3 outlined in detail how the GR might serve as a group fitness maximizing mechanism. Moreover, there are many "be altruistic" instructions in the broader context of the Deuteronomic text mentioned above, including the Ten Commandments. For trait 9, we will find numerous "be groupish" instructions in Deuteronomy. "You may require payment from a foreigner, but you must cancel any debt your brother owes you" (Deuteronomy 15:3). This is the GR applied to insiders only. Incidentally, we also find instructions to "be altruistic" to outsiders as well:

[The Lord] defends the cause of the fatherless and the widow, and loves the alien, giving him food and clothing. And you are to love those who are aliens, for you yourselves were aliens in Egypt. (Deuteronomy, 10:18-19, New International Version)

Moreover, for trait 9, the Deuteronomy 6 passage cited above calls for numerous identity markers that visually show adherence to the meme as a member of the in-group. "Write the meme on your doorposts." Skipping to trait 14, this text does not promise ostracism for those who fail to adhere, but the following context (Deuteronomy Chapter 7), threatens destruction for those who do not adhere. Lastly, regarding the groupish traits 8-9, Wilson (2002: 133-147) provides a detailed description of how Judaism has facilitated altruism and group selection.

For trait 10 and returning to the GR, "do-as-you-would-be-done-to" shows rhyming and alliteration, but can we show that this is more memorable than other ways to render the Golden Rule? This is a good question, and one that we may answer through research. Moreover, we will need to extend such research cross-linguistically to see if other cultural variations of the Golden Rule are easily memorized. Besides doing actual memory tests, we can easily check for alliteration, rhyming, and other phonetic and acoustic factors that may make texts more memorable. However, this version of the Golden rule seems to fit, at least superficially, with trait 10 above.

For trait 11, we should first note that music is a cultural universal -- often associated with religion and moral tradition. Moreover, music relates to language, evolution, and morality inasmuch as *linguistically encoded moral teachings combined with music* might promote or maintain the moral development or behavior of human individuals and groups. Though this topic is large enough for a separate chapter, the basic question here concerns whether we can combine a moral concept with music and if this combination will help that concept to attain a memetic state.

Evidence exists that music combined with language enhances the memory of words (Yalch, 1991; Wallace, 1994; Ho et al., 2003). Moreover, significant traditions exist where groups teach moral concepts set to music showing that morally laden teachings may become more memorable to the individuals and groups who learn them in songs. If this is true, we will want to know if morals set to music can attain a memetic state, can they also positively influence altruistic and moral behavior? In short, this brief mention of music and memes raises more questions than it answers.

However, the questions appear good enough to lead to some potentially fascinating and tangible memetic research, the sort of research that memetics has failed to produce so far.

Trait 12 concerns how the grammatical and lexical rendering of a particular idea might affect its memetic status. As mentioned above, (A) "Do as you would be done to" may be more memorable than (B) "Treat others in the same way that you want them to treat you" because (A) is more succinct. This point also raises a question that transcends the scope of this chapter. Will a concise, succinct, and pithy rendering of an idea cause it to attain a greater memetic status than an idea that is rendered with less precision? This also seems like the kind of question that could drive some interesting research into finding out if subjects might more easily memorize one rendering of the GR than another.

Trait 13 asks if we can reduce the GR to protolanguage. However, we previously saw that the GR requires argument structure (who does what to whom), and thus we cannot reduce its grammatical complexity to protolanguage, which lacks argument structure. Nevertheless, since the GR condenses numerous universal moral rules, we can speculate that it could have emerged out of protolinguistic morality as humans developed argument structure. Initially, early humans could have combined the naming capacity with a degree of creativity, enabling them to label things, events, and behaviors as good or bad. This would be fitness enhancing in two ways. First, verbal labels for right and wrong could form a safer and more cost efficient means for social control than physical enforcement, especially as group size increased (Dunbar, 1996). Second, proto-moral maxims would promote survival behavior and reproductive fitness. For example, the ideas kill-bad, procreate-good, steal-bad, share-good, do not require argument structure but represent ideas that could increase the fitness of individuals and their group, and the GR summarizes these ideas.

Moreover, the phrase "cheat-bad" exemplifies another fitness beneficial idea that can be communicated in protolanguage, which might promote and regulate altruism in small groups. As humans developed increasingly complex language, these moral maxims could be expressed with argument structure, such as, "Don't

cheat on your neighbor." "Do not kill neighbors." "Share food with neighbors." "Make babies with your spouse, not your neighbor's spouse." Moreover, as cognitive abilities developed, humans could logically summarize these ideas in the phrase "do as you would be done to" as a concept (or meme) that would enhance group fitness.

Let us more fully consider the adaptive value of protolinguistic moral memes. First, how would verbal labels for right and wrong more efficiently promote survival behaviors and social control than physical enforcement, especially as group size increased? During development, humans deeply depend on parents for a long time. Moreover, we can safely assume that we linguistically encode and transmit fitness enhancing information that is not already encoded in our genes. That is, the moral memes mentioned here can adaptively guide human development because humans do not fully encode these adaptive ideas (and behaviors) genetically. Obviously, if these putative memes increase fitness enhancing behavior, then they will have an adaptive benefit, and in a large group with many "fitness ignorant" offspring, protolinguistic moral memes would help us efficiently pass on this fitness enhancing information through language.

Nevertheless, though "steal-bad" is a moral idea couched in protolanguage, it is still complex because it assumes some knowledge of social calculus. Stealing is bad because the group punishes me for it, but if I have some basic empathy, then I also sense how my conspecific feels when I steal from him. Thus, we must ask if early humans could have comprehended social calculus or even argument structure without being able to produce it linguistically. Fortunately, research in primate linguistics might prove helpful here. That is, since primates are non-linguistic species that only attain linguistic ability through human instruction, perhaps their ability to comprehend argument structure is separate from (or more basic than) their ability to produce it. This assumes that teaching them language does not also teach them argument structure, which is unlikely.

Regarding ape language, Kako (1999), claims that Kanzi demonstrates some understanding of discrete combinatorics, category-based rules, and argument structure. This is because Kanzi has successfully comprehended sentences where

single word meanings and prior knowledge are insufficient for interpretation and where the roles in a sentence were interchangeable. Thus, Kanzi needed to rely at least partly on syntax to comprehend these utterances. In short, these problems raise the empirical questions regarding what species have the cognitive ability to comprehend argument structure while not possessing language. Thus, if early humans could have comprehended but not produced argument structure, they could have understood protolinguistic moral memes with an implicit argument structure, and these memes could have promoted fitness enhancing behavior.

Trait 14 was discussed above, so we move to Trait 15 (and 7), which concern beliefs and metaphysical factors that memeticists often discuss rather negatively. In this approach, however, we are not concerned whether the metaphysical factors in question are true or false, but whether metaphysical stances can potentially increase the putative memetic state of concepts. For example, let us say that a group believes that if families do not share their food with others the gods will curse them. (Sharing food is group-adaptive GR-type behavior). Groups might recite myths, teaching that non-sharers go insane, and this belief could spread through the group. This metaphysical belief would be thus tied to the altruistic acts. Even though the belief is false, promises and threats tied to *metaphysical belief* in rewards and punishments may affect behavior as much as tangible rewards and punishments. Moreover, this belief tied to a moral rule that exhibits the traits listed here could push this concept into a catalytic and memetic state, and this would give such a meme considerable power over the people who think it.

This scenario suggests research avenues for a lingua-memetic approach for understanding morality that should lead us to ask a number of questions regarding moral concepts that groups connect to metaphysical beliefs. Do these beliefs help people more easily memorize and follow the moral rules tied to them? Does a metaphysical belief in reward or punishment affect the memetic status of the moral rules associated with this belief? That is, will the "share your food meme" more likely spread through a population catalytically if people have strong metaphysical beliefs about it? Do moral memes with metaphysical backing influence altruistic behavior in people? Do acts of sharing increase or decrease if detached from a

metaphysical belief and attached to secular beliefs? In short, instead of vaguely talking about "god memes," we need to tangibly answer these kinds of questions if we are going to actually discover memes that relate to metaphysical beliefs.

This brings us to trait 16, which is not actually a trait that helps an idea attain memetic status, but rather the memetic catalytic state itself. Auger proposes that memes replicate between neuronal nodes within brains, and this is helpful in the subsidiary sense that concepts actually replicate *within* minds. However, this does not imply that putative memes can replicate between brains, and if social memetic replication remains unconfirmed, then memetics will fail to serve as a science of culture and the minds that make it. Thus, the key question concerns how these 16 traits might help create memes that socially replicate between brains through language in degrees independent of human intention.

Human linguistic ability, to a degree, exists instinctively in human minds, and is thus not a product of our intentions, but of our biological endowment. In this way, language happens to us as much as we make it happen because our biology genetically programs linguistic ability into us. For memes, we want to find out if culture can memetically program a particular kind of idea into us that will spread in degrees beyond our intentional control. One possible approach denies intentional agency all together; however, this may be untenable and may never become a methodological position based on an empirical rationale. This is like extreme versions of selfish gene theory that presume selfish selectionism, preempting the possibility of multi-level selection before empirical inquiry begins. That is, we cannot empirically prove that memes move above human intention by preemptively denying the existence of intention in the first place. Keeping intentionality in the picture leaves us with a messy state of affairs that will not satisfy those who want to turn sociology, linguistics, and memetics into hard science. However, it remains an overwhelming possibility that this is the reality we have to deal with.

In this messy situation, these 16 traits might increase the probability of social memetic replication, and research into these traits might reveal indirect evidence that ideas can partly possess us. Thus, these 16 traits may give us ways to research *how* ideas might attain a memetic state, and this relates to how language helps

develop, maintain, facilitate, and transmit human altruism and morality. This is because moral memes (morals in a memetic state) will affect us more powerfully than non-memetic moral ideas. Besides this, proto-moral-memes, couched in human protolanguage could have powerfully extended early human morality. Moreover, if proto-morality and protolanguage helped increase the fitness of early human groups, this could have also helped their language to develop and evolve. This will be hard to prove; nevertheless, the above traits suggest how we can investigate the fitness of memes and how they might catalytically replicate between minds. And if we can demonstrate the catalytic replication of moral memes, this would be a firm but humble step in the right direction for the field of memetics.

CHAPTER 8

CONCLUSIONS

8.1 Summary

This thesis has covered numerous specific questions related to language and morality. Though its focus may appear to be broad, it is actually firmly constrained by one key question. *Given the insights of evolutionary theory, how does language influence human altruism and morality?*

In Chapter 2, this question directed us to consider various functions of morality that language enables or extends. These functions related to accessing and enforcing moral norms, sharing moral representations, and facilitating cooperation. Moreover, we considered numerous parallels between linguistic and moral ontogeny in humans, looking at the relationship between language, empathy, altruism, and morality. We also briefly reviewed altruism in animals that cannot use language to help us understand how language affects animals and humans who can use language. This chapter therefore subsumed and reviewed numerous strands of research, relating them all to our key question and strongly demonstrating the essential role that language plays to enable and influence human altruism and morality.

In Chapter 3, we looked at the influence of language on morality and altruism, concerning groups and group selection, and we noted that language can help a group to function as an organism. Moreover, language helps groups put various pressures on individuals that limit cheating and impose social controls, thus enabling them to decrease selfish behavior and increase altruistic behavior that benefits the group. This suggests that the group that can better facilitate altruism

among its members may have a fitness edge over groups that cannot facilitate as much altruism. Instead of emphasizing biologically based altruism that might become genetically coded and transmitted, Chapter 3 emphasized culturally based altruism, stressing the role that a linguistically based moral culture might play in facilitating altruism in groups. Central to this discussion stands the concept of virtual relatedness (*vr*) and the idea that Golden Rule (*g*) type principles facilitate *vr*. Virtual relatedness measures behavioral relatedness, not biological relatedness and suggests that behaviors distinct from inclusive fitness and reciprocal altruism can emerge in groups who use a *g* code that promotes altruism. A *g* code based in language can help us recategorize in-group non-kin as if they were kin and therefore behave towards them as if they were kin. The concept of *vr* emerges as a key contribution of this thesis; however, if we only consider *vr* in the context of in-groups, this raises the question about the dark side of group selection and whether humans can show real altruism to out-group non-kin.

Though proponents of selfish selection generally reject group selection (GS) for reasons mentioned in Chapter 3, GS still meshes with selfish selection theory because it simply elevates selfishness up to the level of between-group competition. Thus, Chapter 4 discusses how *vr* might help humans produce real altruism for out-group non-kin. Specifically, it suggests that we can apply the principle of virtual relatedness to out-group non-kin, thus categorizing non-kin outsiders as if they were insider kin. Moreover, Chapter 4 focused on the issues of reductionism and moral non-realism and claimed that these factors influence current thinkers to mistakenly categorize apparently real altruistic behavior as maladaptive or as manipulated by selfish agents. It was thus suggested that we employ an *a posteriori* critical realism when we look at the problem of real altruism; critical realism will allow us to see the world as a multi-layered stratified reality that requires us to employ different explanatory tools depending on the nature of the strata we study.

Thus, this chapter claims that humans can use language based *g* codes to facilitate *vr* towards non-kin outsiders and it cites examples, such as holocaust rescue, to make its point. Though we saw that humans can use language to promote egoism,

we also noted that language helps us facilitate altruistic ability and plays an essential role in helping human beings develop genuinely altruistic behaviors and personalities. The key contribution of this chapter comes from how it incorporates current research on evolution and altruism and yet still reasonably challenges the thinking that preemptively denies the possibility of real altruism occurring in nature.

Having looked at the various ways that language can influence altruistic behavior in humans, Chapter 5 considered the actual properties of language that enable or enhance morality. Thus, we looked at the relationship between the human moral capacity and universal grammar; moreover, we looked at features common to all human languages that can influence human morality, including: recursion, displacement, creativity, naming ability, semantic primitives, stimulus freedom, cultural transmission, and categorization. With its novel summation of how linguistic properties enhance morality, this chapter stands as another key contribution of this thesis, and it clearly demonstrates how these properties of language greatly extend moral thinking and behaving.

After looking at how specific properties of language influence human altruistic thinking and behaving, Chapter 6 turned our attention to the related question of how differences between languages might affect differences in moral thinking and behaving across cultures. We thus looked at the question of linguistic relativity as it relates to how differences between English and Japanese might help create very subtle but still significant differences in moral thinking in bilingual speakers of these languages. Though the results did not attain statistical significance, they were still promising enough to encourage future research on this topic. Most interestingly we saw the potential influence of language and gender on female speakers of Japanese and English. Namely we noted that using English might cause them to face conflict and communicate more directly, and using Japanese might cause them to avoid conflict and communicate more indirectly.

In addition, we noticed that even the best possible translation pairs of some Japanese and English words only approximate the same meanings in both languages, thus showing that language may underdetermine our understanding of

moral concepts across cultures. Though the results remain inconclusive, this chapter still stands as a valuable contribution of this thesis because it attempts novel research in linguistic relativity, focusing on how linguistic differences might facilitate subtle differences in moral thinking across cultures. It also is a valuable contribution because of the promise it holds for future research.

Chapter 7 looked at the up to now unproductive field of memetics as a way of studying cultural transmission. It suggested a linguistic approach for rectifying problems in memetics in order to see if we can actually find ways that culture and moral values might replicate independently or semi-independently of human intention. Though we still have no evidence for real memes or an independent replicator dynamic for memes, this chapter outlined an albeit humble but also realistic agenda for lingua-memetics by suggesting 16 factors that might cause ideas couched in language to attain a memetic state. This relates to our main question on how language influences morality because moral systems put strong pressures on individuals and groups to replicate their moral ideas or memes through language. This chapter holds 2 key contributions. (1) It novelly combines previously suggested memetic factors with altruistic, moral, and linguistic factors that might increase the memetic state of some ideas. (2) It suggests a realistic and limited version of memetics that is linguistically oriented and that has potential for future research.

8.2 Future Research

The question of how language influences morality has proven to be a productive one. Not only has it provided a rich soil from which this thesis has developed, it also holds promise for research. First, we can further explore the relationship and parallels between moral and linguistic ontogeny in small children. For example, we can look more closely at the stages of linguistic development in children, the grammar and lexis they learn, and we can see how they communicate with this language concerning topics related to altruism and morality. We can also look more closely at how language can facilitate empathy (or opposite emotions), and

we can search for objective ways to linguistically elicit and measure these empathic and non-empathic responses in research subjects.

Second we can employ the concept of virtual relatedness as a means for evaluating behavior that appears to be genuinely altruistic. For example, concerning situations where strong enmity exists between rival groups, we can look for counter examples of individuals from these feuding groups who actually behave altruistically to each other. Thus, we can develop a measure for evaluating why they behave this way and see if it coheres with the concept of virtual relatedness. Third, future research also can focus on how linguistic differences might affect subtle differences in moral perception across cultures. The results to the questionnaire in Chapter 6 showed us a specific course to take concerning gender, linguistic, and behavioral differences that relate to harmonious versus directness language when comparing English and Japanese. We can develop a new survey focusing on this specific topic and attempt a larger study than the one conducted for this thesis. Lastly, concerning memetics, we can begin to look for ways that putative memes might replicate somewhat independently of human intention. For example we can test whether ideas that are bonded to the 16 factors mentioned in Chapter 7 become more memorable than ideas that are less bonded to these factors. With all these questions in mind, it seems clear that the relation between evolution, language, and morality promises to be a productive research topic that is of much theoretical interest for scholars as well as serious practical value for human welfare. Therefore, it surely merits our best scholarship and most serious attention.

APPENDIX A & B

BILINGUAL MORAL DILEMMA SURVEYS

The following surveys were used in the experiment described in Chapter 6.

A.1 English Survey Page 1

Real people have experienced all these dilemmas. Imagine you are in their shoes. Choose one solution per dilemma. Try to respond to each situation as you would in real life. To help maintain your privacy, please type your subject code here: <input type="text"/>	A	B
1. 1. You find 10,000 yen (\$100 US) on the street, and you are alone and no one can see you. Would you keep the money (A) or turn it into a police station (B)?	<input type="checkbox"/>	<input type="checkbox"/>
2. 2. You were in a small car accident. Your insurance agent suggests you lie about the details to keep your premiums down. He tells you this is common practice. Would you follow his advice for the money (A) or decline the advice to maintain integrity (B)?	<input type="checkbox"/>	<input type="checkbox"/>
3. 3. Your co-worker always comes late to work and dresses poorly. You know the boss is unhappy and is considering firing your colleague. Would you warn her (A) or avoid confrontation (B)?	<input type="checkbox"/>	<input type="checkbox"/>
4. 4. You are walking down the sidewalk. A man bumps into you and drops his packages. Would you apologise and help him with the packages (A) or simply help him with his packages (B)?	<input type="checkbox"/>	<input type="checkbox"/>
5. 5. You work for a famous company in your country. A head-hunter from a foreign and equally famous competitor calls and offers you a similar job with better pay. Would you (A) accept their offer or (B) turn them down?	<input type="checkbox"/>	<input type="checkbox"/>
6. 6. A woman co-worker of equal status is having a busy day and asks you to make 120 copies of some documents. On an average day, would you copy them for her (A), or would you think of some polite excuse to decline (B)?	<input type="checkbox"/>	<input type="checkbox"/>
7. 7. At work it is against the rules to do personal email on company computers. But almost all employees do personal email at work. In this situation, would you do some personal email at work (A) or follow the rule strictly (B)?	<input type="checkbox"/>	<input type="checkbox"/>
8. 8. A volunteer comes to your door from a well known charity. He politely tries to convince you--on logical grounds--to donate money, saying that small gifts really make a difference. Would you likely be convinced by his logic (A) or turned off by it (B)?	<input type="checkbox"/>	<input type="checkbox"/>
9. 9. You are a manager. One middle-aged employee regularly fails to come to work, giving no reason, no excuse. Would you tell him to change because (A) this behaviour is stealing or (B) because his co-workers need his help?	<input type="checkbox"/>	<input type="checkbox"/>
10. 10. During war, a sniper kills 5 of your buddies. You catch him and have the opportunity to kill him, and since there are no witnesses, you can kill him without getting caught-even though it is against international law. Would you (A) avenge your friends or (B) hand him over as a prisoner?	<input type="checkbox"/>	<input type="checkbox"/>
11. 11. Your close friend (female, unmarried) is having a love affair with a married man whom you do not know. Would you (A) tell your friend to stop the affair because you think it is wrong, or (B) not say anything and maintain harmony in your relationship?	<input type="checkbox"/>	<input type="checkbox"/>
12. 12. Your friend invites you to a party, but you do not want to go. Would you tell him directly that you are not interested in going (A), or would you think of some polite or indirect excuse for not going (B)?	<input type="checkbox"/>	<input type="checkbox"/>
13. 13. Your company pollutes a water supply. Unless you publicise this, about 25,000 people will die in 50 years. Your boss, who has Mafia	<input type="checkbox"/>	<input type="checkbox"/>

links, offers you \$1,000,000 to keep quiet, saying he'll kill your family if you don't. Will you (A) keep quiet to protect your family or (B) speak out to save the city?		
14. 14. Your friend feels considerable guilt because he believes he sinned against his wife by having a love affair, which is over now. Would you (A) comfort him and tell him not to worry because he is sorry and his wife does not need to know, or (B) tell him he should confess and apologize to his wife to save her honor?	A <input type="checkbox"/>	B <input type="checkbox"/>
15. 15. As a teacher, some students fail your exams. The boss tells you to pass them, for they give your school big income. You disagree, but your boss insists, blaming you (falsely) for teaching them poorly. Would you (A) accept his authority and fix the exam results or (B) challenge his instruction?	A <input type="checkbox"/>	B <input type="checkbox"/>

Background Information**Gender:**

- ☐ Male.
☐ Female.

Age:

- ☐ 18-29.
☐ 30-39.
☐ 40-59.
☐ 60+.

Education:

- ☐ High School.
☐ University.
☐ Graduate.
☐ Post Grad

What is your first language?

- ☐ English.
☐ Japanese.

Other?**What is the level of your English reading ability?**

- ☐ Graduate Level.
☐ Undergraduate Level
☐ High School Level.
☐ Junior High School Level.

Press the "reset" button if you want to start over.

Press the "submit" button when you are finished.

Click here to go to the [exit page](#).

B.1 Japanese Survey Page 1

ここに挙げたジレンマは実在の人物が実際に経験したものです。自分が同じような経験をしたら、どのようにそのジレンマを解決するか1つ選んでください。記名の必要はありません。実際にその状況に遭遇したと想定して答えてください。なお、裏面にもご記入下さい。あなたのプライバシーを守るため、あなたのサブジェクトコードをここにタイプしてください <input type="text"/>	A	B
1. 道ばたで一万円を見つけました。そこにいるのはあなただけで、誰もあなたをみていません。(A)あなたはそのお金を取りますか?(B)それとも、交番に届けますか?	A	B
2. あなたは小さな交通事故に遭いました。保険会社の保険員はあなたの保険利率を保つため、事故について嘘をつくように提言します。このようなことは誰もがすることだと保険員はいいいます。あなたは(A)お金のために彼のアドバイスのに従いますか、(B)それとも、高潔さを保つためそのアドバイスを断りますか?	A	B
3. あなたの仕事上の仲間は遅刻をし、身だしなみもよくありません。あなたの上司はそのことを快く思わず、あなたの同僚を解雇しようとしていることをあなたは知っています。あなたは(A)彼女にこのことを告げますか、(B)それとも直接に彼女と話すことは避けますか?	A	B
4. あなたは歩道を歩いています。ある男性がぶつかってきて、彼が持っていたものを落とします。(A)あなたは謝り、荷物を拾うの手伝いますか、(B)それとも、ただ荷物を拾うことを手伝いますか?	A	B
5. あなたは自国の有名企業で働いています。あなたの会社とほぼ同等に有名な外国のライバル社のヘッドハンターが電話をしてきて、同じ職種の仕事をさらに良い給料でオファーします。あなたはこのオファーを(A)受けますか?または、(B)断りますか?	A	B
6. あなたと同じ地位をもった女性の同僚がとても忙しく働いているある日、書類を120枚コピーしてほしいとあなたに頼んできました。あなたの仕事は平常通りです。あなたは彼女のために(A)コピーをしますか、(B)それとも、失礼にならないように言い訳を考えて断りますか?	A	B
7. あなたの職場では、会社のコンピューターで私用のEメールをすることは禁じられています。しかし、ほとんどの者は職場で私用のEメールをしています。この状況で、(A)あなたは職場でEメールを私用に使用しますか、(B)それとも、厳密に会社のルールに従いますか?	A	B
8. ある有名なチャリティーのボランティア員があなたの家に来ました。彼は丁寧に、小さな寄付をすることが大きな変化をもたらすと、論理的にあなたを説得します。あなたは、(A)論理的な説明に説得されるでしょうか、(B)それとも、論理的な説明によって逆に断りたくなるでしょうか?	A	B
9. あなたは管理職についています。ある中年の社員は理由もなく、また断りを入れることもなく、よく仕事を休みます。あなたは彼の行為を改めるためどのようにしますか?(A)この行為は盗みであるから彼の行為は改めなくてはならない。(B)彼の同僚が彼の助けを必要としているから、その行為を改めなくてはならない。	A	B

10. 戦争中、ある狙撃手があなたの仲間5人を殺しました。あなたはその狙撃手を捕まえ、彼を殺すチャンスを得ます。国際法に違反はしますが、誰も見ていないので捕まえられることなくこの狙撃手を殺すことができます。あなたは(A)あなたの味方の仇をうちますか、(B)それとも、捕虜として渡しますか？	A	B
11. あなたの親しい友人(女性、独身)が、あなたの知らない既婚の男性と不倫関係にあります。あなたは(A)あなたの友人に不倫はまちがったことであるからやめるように、と言いますか？(B)それとも、友人との関係にヒビが入らないよう何も言いませんか？	A	B
12. あなたの友人があなたをパーティーに誘っていますが、あなたはあまり乗り気ではありません。あなたは(A)彼にストレートに興味がないと言いますか？(B)それとも、丁寧に断りますか？	A	B
13. あなたの会社は水質汚染をしています。あなたがこのことを明らかにしないと、50年で約25,000人が死んでしまいます。あなたの上司はマフィアとつながりがあり、黙認すれば1億2千万円を、もし公表すればあなたの家族を殺すと言います。あなたは(A)家族を守るために黙っていますか、それとも(B)町を救うために事実の公表をしますか？	A	B
14. あなたの友人は、すでに終わっている1年ほど続いた不倫関係を持ったことで彼の妻に対して罪を犯したと信じ、彼の妻に対してかなりの罪悪感があります。あなたは(A)彼を慰め、彼は反省しているし、彼女は何も知る必要はないと言いますか、それとも(B)全てを告白し謝罪することで彼女に敬意を払うと言いますか？	A	B
15. 教師であるあなたは何人かの学生を試験で不合格にしました。あなたの上司は、彼らは学校にとって大きな収入源であるから彼らをパスさせるようにいいます。あなたはそれに反対しますが、あなたの上司はあなたの教え方が悪いとあなたを(不当に)責めながら、どうしても学生をパスさせるようにいいます。あなたは(A)上司の権威を認めて試験結果を修正しますか、それとも(B)上司の指示に挑みますか？	A	B

Background Information

性別:

☐ 男☐ 女

年齢:

☐ 18歳~29歳☐ 30歳~39歳☐ 40歳~59歳☐ 60歳以上

学歴:

☐ 高卒☐ 大卒☐ 修士☐ 博士

あなたの第1言語は何ですか？

☐ 英語

☐ 日本語

他の？

あなたの読解力は何のレベルですか？

☐ 大学院

☐ 大学

☐ 高校

☐ 中学校

もう一度やりなおす場合は"reset"をクリックしてください。

Reset

終了した方は"submit"をクリックしてください。

Submit

出口へ、[ここ](#)をクリックしてください。

REFERENCES

- Aitchison, J. (1999). *Linguistics: An introduction*. London: Hodder & Stoughton.
- Anderson, S., Bechara, A., Damasio, H., Tranel, D., & Damasio, A. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience*, 2:11, 1032-37.
- Arnold, K. & Whiten, A. (2001). Post-conflict behaviour of wild chimpanzees (pan troglodytes schweinfurthii) in the Budongo Forest, Uganda. *Behaviour*, 138, 649-90.
- Aunger, R. (2002). *The electric meme: a new theory of how we think*. New York: Free Press.
- Aunger, R. (Ed.). (2000). *Darwinizing culture: The status of memetics as science*. London: Oxford University Press.
- Balazs, A. (2004). *Biosystems*, 73(1), 1-11.
- Balkin, J. M. (2003). *Cultural Software: A Theory of Ideology*. Yale University Press.
- Batson, D. C. (2002). Addressing the altruism question experimentally. In S. G. Post, L. G. Underwood, J. P. Schloss, & W. B. Hurlbut (Eds.), *Altruism and altruistic love*. London, UK: Oxford University Press.
- Beck, A. T., & Freeman, E. D. (1990). *Cognitive Therapy of Personality Disorders*. London: The Guilford Press.
- Beck, R. & Fernandez, E. (1998). Cognitive-behavioral therapy in the treatment of anger: a meta-analysis. *Cognitive Therapy and Research*, 22:1, 63-74.
- Becker, H. (1956). *Man in reciprocity*. New York: Praeger.
- Bekoff, M. (2001). Social play behavior: Cooperation, fairness, trust, and the evolution of morality. *Journal of Consciousness Studies*, 8(2), 81-90.
- Bekoff, M. (2002). Virtuous Nature. *New Scientist*, 176, 34.

- Belpaeme, T. (2002). *Factors influencing the origins of colour categories*. Doctoral Thesis: Artificial Intelligence Laboratory Vrije Universiteit Brussel.
- Bernstein, I.S. (2000). The Law of Parsimony Prevails: Missing Premises Allow Any Conclusion. In L. Katz (Ed.), *Evolutionary Origins of Morality: Cross-Disciplinary Perspectives*. UK: Thorverton Imprint Academic.
- Bhaskar, R. (1987). *Scientific Realism and Human Emancipation*. Verso Books.
- Bhaskar, R. (1997a). On the ontological status of ideas. *Journal for the Theory of Social Behavior*, 27:2/3, 139-47.
- Bhaskar, R. (1997b). *A Realist Theory of Science (Verso Classics, 9)*. Verso.
- Bhatia, S. (2000). Language socialization and the construction of socio-moral meanings. *Journal of moral education*, 29:2, 149-66.
- Bickerton, D. (1981). *The roots of language*. Ann Arbor, Mich.: Karoma.
- Bickerton, D. (1996). *Language and Human Behavior (The Jessie and John Danz Lectures)*. University of Washington Press.
- Blackmore, S. (1999). *The meme machine*. London, UK: Oxford University Press.
- Blackmore, S. (2000). The power of memes. *Scientific American*, 283, No 4, 52-61.
- Blair, R. & Cipolotti, L. (2000). Impaired social response reversal: A case of 'acquired sociopathy'. *Brain*, 123, 1122-41.
- Bloch, M. (2000). A well-disposed social anthropologist's problems with memes. In R. Aunger (Ed.), *Darwinizing culture: The status of memetics as science*. London: Oxford University Press.
- Boehm, C. (2000). Conflict and the Evolution of Social Control. In L. Katz (Ed.), *Evolutionary Origins of Morality: Cross-Disciplinary Perspectives*. UK: Thorverton Imprint Academic.
- Boom, C. T., & Buckingham, J. (1986). *Tramp for the Lord*. Jove Pubns.
- Boroditsky, L. (2001). Does language shape thought? English and Mandarin speakers' conceptions of time. *Cognitive Psychology*, 43:1, 1-22.
- Boroditsky, L. (2002). Interview on National Public Radio.
<http://cognition.mit.edu/press/interview-high-quality.html>
- Boroditsky, L. (2003). Linguistic relativity. In *Encyclopedia of Cognitive Science*. London: MacMillan Press.
- Boyd, R., & Richerson, P. (2000). Memes: Universal Acid, or a Better Mouse Trap? In R. Aunger (Ed.), *Darwinizing culture: The status of memetics as science*.

London: Oxford University Press.

Boyd, R., & Richerson, P. J. (1988). *Culture and the Evolutionary Process*. University Of Chicago Press.

Brodie, R. (2004). *Virus of the Mind: The New Science of the Meme*. Integral Press.

Brosnan, S. F. & De Waal, F. B. (2003). Monkeys reject unequal pay. *Nature*, 425(6955), 297-99.

Brower, M. & Price, B. (2001). Neuropsychiatry of frontal lobe dysfunction in violent and criminal behaviour: a critical review. *Advances in neuropsychiatry*, 71, 720-26.

Brown, D. E. (1991). *Human universals*. New York: McGraw-Hill.

Brown, W., & Paul, L. (2002). Broken light: Wisdom and the disconnected brain. *Science and Spirit*. Retrieved February 13, 2006, http://www.science-spirit.org/articles/articledetail.cfm?article_id=124

Buber, M. (1971). *I And Thou*. Free Press.

Calvin, W., & Bickerton, D. (2000). *Lingua ex Machina: Reconciling Darwin and Chomsky with the Human Brain*. Cambridge, MA: MIT Press.

Cangelosi, A. H., S. (2001). The Adaptive Advantage of Symbolic Theft Over Sensorimotor Toil: Grounding Language in Perceptual Categories. *Evolution of Communication*, 4(1), 117-42.

Chomsky, N. (1959). A review of B.F. Skinner's Verbal Behavior. *Language*, 35(1), 26-58.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, Mass: MIT Press.

Christiansen, M., & Kirby, S. (2003). *Language Evolution (Studies in the Evolution of Language)*. Oxford University Press.

Cloak, F. (1975). Is a cultural ethology possible? *Human Ecology*, 3(3), 161-82.

Colin, A. (2004). Teleological Notions in Biology. *The Stanford Encyclopedia of Philosophy*.

Condon, J. (1984). *With respect for the Japanese*. Yarmouth, ME: Intercultural Press.

Conte, R. (2000). Memes Through (Social) Minds. In R. Aunger (Ed.), *Darwinizing culture: The status of memetics as science*. London: Oxford University Press.

Corrie Ten Boom: History. Retrieved February 8, 2006, from <http://www.corrietenboom.com/history.htm>

- Crystal, D. (1992). *Introducing linguistics*. London: Penguin.
- Crystal, D. (1997). *The Cambridge encyclopedia of language*. New York: Cambridge University Press.
- Daniel, S. (2005). Physicalism. *The Stanford Encyclopedia of Philosophy*.
- Darwin, C. (1859). *The origin of the species*. Retrieved February 13, 2006, <http://www.literature.org/authors/darwin-charles/the-origin-of-species/index.html>
- Darwin, C. (1882). *The descent of man, and selection in relation to sex*. Retrieved February 13, 2006, <http://www.literature.org/authors/darwin-charles/>
- Dawkins, R. (1976). *The Selfish Gene*. Oxford, UK: Oxford University Press.
- Dawkins, R. (1982). *The extended phenotype: the long reach of the gene*. Oxford, UK: Oxford University Press.
- de Waal, F. (1982). *Chimpanzee politics: Power and sex among apes*. London: Counterpoint.
- de Waal, F. (1996). *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Cambridge, MA: Harvard University Press.
- de Waal, F. (2002). Personal communication.
- Deacon, T. (1997). *The symbolic species: The co-evolution of language and the human brain*. London, UK: Penguin Books.
- DeMoivre, and the Evolutionary Ideas of Darwin, Düsing, Fisher, Williams, and Hamilton. <http://philosophy.wisc.edu/sober/papers.htm>
- Dennett, D. (1995). *Darwin's dangerous idea: Evolution and the meanings of life*. UK: Harmondsworth.
- Dessalles, J. L. (1998). Altruism, Status, and the Origin of Relevance. In J. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the Evolution of Language - Social and Cognitive Bases*. London: Cambridge University Press.
- Dicks, L. (2000). All for One. *New Scientist*, 167, 30.
- Dolan, R. (1999). On the neurobiology of morals. *Nature Neuroscience*, 2:11, 927-28.
- Dover, G. (2001). Anti-Dawkins. In H. Rose, & S. Rose (Eds.), *Alas poor Darwin: Arguments against evolutionary psychology*. New York: Random House.
- Dunbar, R. (1996). *Grooming, Gossip and the Evolution of Language*. London: Faber and Faber.

- Dunn, J. (1987). The beginnings of moral understanding: Development in the second year. In J. Kagan, & S. Lamb (Eds.), *The emergence of morality in young children*. Chicago: University of Chicago Press.
- Durham, W. (1991). *Coevolution: Genes, culture, and human diversity*. Stanford, CA: Stanford University Press.
- Edwards, P. (Ed.). (1967). *The encyclopedia of philosophy*. New York: Macmillan Publishing Company.
- Ellis, R. (1997). *Second language acquisition*. London: Oxford University Press.
- Endler, J. A. (1986). *Natural Selection in the Wild*. Princeton University Press.
- Enright, S. (1997). Fortnightly review: Cognitive behaviour therapy-clinical applications. *British Medical Journal*, 315:7119, 1376-77.
- Ervin-Tripp, S. (2001). Personal communication.
- Forgas, J. & Ciarrochi, J. (2002). On managing moods: Evidence for the role of homeostatic cognitive strategies in affect regulation. *Personality and social psychology bulletin*, 28:3, 336-45.
- Forgas, J. (2001). Affect and information processing strategies. In J. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition*. London: Cambridge University Press.
- Fouts, R., & Mills, S. T. (1998). *Next of Kin: My Conversations with Chimpanzees*. Harper Paperbacks.
- Galasinski, D. (2000). *The Language of Deception: A Discourse Analytical Study*. Sage Publications.
- Glass, J. M. (1999). *"Life Unworthy of Life": Racial Phobia and Mass Murder in Hitler's Germany*. Basic Books.
- Goddard, C. (1997). The universal syntax of semantic primitives. *Language Sciences*, 19(3), 197-207.
- Goddard, C. (2003). Whorf meets Wierzbicka: variation and universals in language and thinking. *Language Sciences*, 25(4), 393-432.
- Goddard, C., & Wierzbicka, A. (1994). *Semantic and Lexical Universals: Theory and Empirical Findings*. John Benjamins Publishing Company.
- Goddard, C., Wierzbicka, A., & O'Neill, K. J. (2002). *Meaning and Universal Grammar: Theory and Empirical Findings*. John Benjamins Publishing Co.
- Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology, General*, 123, 178-200.

- Goodenough, U. & Deacon, T. (2003). From biology to consciousness to morality. *Zygon*, 38:4, 801-19.
- Gould, S. (2001). More things in heaven and earth. In H. Rose, & S. Rose (Eds.), *Alas poor Darwin: Arguments against evolutionary psychology*. London, UK: Random House.
- Gratzer, W. (2000). *The undergrowth of science: delusion, self-deception, and human frailty*. London: Oxford University Press.
- Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. L. Morgan (Eds.), *Speech acts*. New York: Academic Press.
- Gumperz, J., & Levinson, S. (Eds.). (1996). *Rethinking linguistic relativity*. London, UK: Cambridge University Press.
- Hall, E. T., & Hall, E. T. (1977). *Beyond Culture*. Anchor.
- Hamilton, W. (1964). The genetical evolution of social behavior, I and II. *Journal of Theoretical Biology*, 7, 1-52.
- Hamilton, W. (1971). The Genetical Evolution of Social Behavior I-II. In G. C. Williams (Ed.), *Group Selection*. New York: Aldine • Atherton.
- Hamilton, W. D. (1967). Extraordinary sex ratios. *Science*, 156, 477-88.
- Hanfling, O. (2003). Learning about right and wrong: Ethics and Language. *Philosophy*, 78, 25-41.
- Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Anim Behav*, 61(1), 139-51.
- Harms, W. (2000). Adaptation and moral realism. *Biology and Philosophy*, 15(5), 699-714.
- Harnad, S. (1987). Category Induction and Representation. In S. Harnad (Ed.), *Categorical Perception: The Groundwork of Cognition*. (Chapter 18). New York: Cambridge University Press.
- Harnad, S. (2003). Categorical Perception. Retrieved February 13, 2006, <http://www.ecs.soton.ac.uk/~Eharnad/Temp/catperc.html>
- Harnden-Warwick, D. (1997). Psychological Realism, Morality, and Chimpanzees. *Zygon*, 32(1), 29.
- Hartung, J. (1995). *A Light Unto the Nations: Judeo-Christianity, Morality, and Group Selection*. Human Behavior and Evolution Society Conference June 28-July 2, 1995 at the University of California, Santa Barbara.
- Hauser, M. D. (1992). Costs of deception: cheaters are punished in rhesus monkeys

(*Macaca mulatta*). *Proc Natl Acad Sci U S A*, 89(24), 12137-39.

- Hauser, M., Chomsky, N., & Fitch, T. (2002). The faculty of language: What is it, Who has it, and How did it evolve? *Science*, 298(22), 1569-79.
- Ho, Y.-C., Cheung, M.-C., & Chan, A., S. (2003). Music Training Improves Verbal but Not Visual Memory: Cross-Sectional and Longitudinal Explorations in Children. *Neuropsychology*, 17(3), 439-50.
- Hodgson, G. (2001). Darwinism and Evolutionary Economics. In J. Laurent, J. Nightingale, & J. Laurent (pp. 87-118). Edward Elgar Publishing.
- Hull, D. (2000). Taking memetics seriously: memetics will be what we make it. In R. Auger (Ed.), *Darwinizing culture: The status of memetics as science*. London: Oxford University Press.
- Hunter, C. (2001). *Darwin's God: Evolution and the Problem of Evil*. Grand Rapids, Michigan: Brazos Press.
- Hurford, J. (1998). The emergence of syntax. In J. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language*. Cambridge, UK: Cambridge University Press.
- Hurford, J., Studdert-Kennedy, M., & Knight, C. (Eds.). (1998). *Approaches to the evolution of language*. Cambridge, UK: Cambridge University Press.
- Huxley, T. H. (1893). Evolution and Ethics: The Romanes Lecture. Retrieved February 7, 2006, from <http://cs.clarku.edu/huxley/CE9/E-E.html>
- Inoue, S. (1998). The Japanese Language which Prefers Ambiguity vs. the English Language which Prefers the Battle of Words. *Research Journal of Tokyo University of Information Science*, 2:2, 91-97.
- Irwin, L. (1997). The Web Site for Critical Realism (WSCR): Ethical naturalism and moral realism. Retrieved February 13, 2006, <http://www.raggedclaws.com/criticalrealism>
- Jackendoff, R. (2002). Foundations of language: Brain, meaning, Grammar, Evolution. New York: New York: Oxford University Press.
- Kagan, J. (2002). Morality, altruism, and love. In S. G. Post, L. G. Underwood, J. P. Schloss, & W. B. Hurlbut (Eds.), *Altruism and altruistic love*. London, UK: Oxford University Press.
- Kako, E. (1999). Elements of syntax in the systems of three language-trained animals. *Animal Learning and Behavior*, 27:1, 1-14.
- Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: evidence for "fast mapping". *Science*, 304(5677), 1682-83.

- Katz, L. (Ed.). (2000). *Evolutionary Origins of Morality: Cross-Disciplinary Perspectives*. UK: Thorverton: Imprint Academic.
- Kirby, S. (2000). Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In C. Knight, M. Studdert-Kennedy, & J. Hurford (Eds.), *The evolutionary emergence of language: social function and the origins of linguistic form*. Cambridge, UK: Cambridge University Press.
- Knauff, B. M. (2000). Symbols, Sex, and Sociality in the Evolution of Human Morality. In L. Katz (Ed.), *Evolutionary Origins of Morality: Cross-Disciplinary Perspectives*. UK: Thorverton: Imprint Academic.
- Knight, C., Studdert-Kennedy, M., & Hurford, J. (Eds.). (2000). *The evolutionary emergence of language: social function and the origins of linguistic form*. Cambridge, UK: Cambridge University Press.
- Kuper, A. (2000). If memes are the answer, what is the question?. In R. Aunger (Ed.), *Darwinizing culture: The status of memetics as science*. London: Oxford University Press.
- Laland, K. N., Odling-Smee, F. J., & Feldman, M. W. (2000). Group Selection: A Niche Construction Perspective. In L. Katz (Ed.), *Evolutionary Origins of Morality: Cross-Disciplinary Perspectives*. UK: Thorverton; Imprint Academic.
- Lane, H. (1965). The motor theory of speech perception: A critical review. *Psychological Review*, 72, 275 - 309.
- Larson, D. (1984). *Guidelines for barefoot language learning*. St. Paul, Minnesota: CMS Publishing.
- Lawrence, D. H. (1950). Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation. *Journal of Experimental Psychology*, 40, 175 - 188.
- Lemonick, M. (1997). Young, single, and out of control. *Time*, 68.
- Lenneberg, E. (1967). *Biological foundations of language*. New York: John Wiley.
- LePage, R. B., & Tabouret-Keller, A. (1985). *Acts of Identity: Creole-based Approaches to Language and Ethnicity*. New York: Cambridge University Press.
- Lewis, C. S. (2003). *Perelandra (Space Trilogy (Paperback))*. Scribner.
- Lewontin, R. (1993). *Biology As Ideology: The Doctrine of DNA*. New York: Harper Perennial.
- Lewontin, R. (1998). Survival of the nicest? *New York Review of Books*. Retrieved

February 13, 2006, http://www.nybooks.com/articles/article-preview?article_id=707.

- Lucy, J. (1992). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. London, UK: Cambridge University Press.
- Lutgendorf, P. (2004). The Oral Tradition and the Many "Ramayanas." Retrieved February 13, 2006, <http://www.maxwell.syr.edu/maxpages/special/ramayana/Intro.htm>
- Lykken, D. (2000). The causes and costs of crime and a controversial cure. *Journal of Personality*, 68:3, 559-605.
- Lynch, A. (1996). *Thought contagion: How belief spreads through society*. New York: Basic Books.
- Masserman, J., Wechkin, S., & Terris, W. (1964). "Altruistic" behavior in rhesus monkeys. *American Journal of Psychiatry*, 121, 584-85.
- Maynard-Smith, J., & Szathmary, E. (1999). *The origins of life*. London: Oxford University Press.
- Mayr, E. (1997). The objects of selection. *Proceedings of the national academy of science, USA*, 94, 2091-94.
- McGrath, A. (2004). *Dawkins' God: Genes, Memes, And The Meaning Of Life*. Blackwell Publishers.
- McGrath, A. E. (2002). *A Scientific Theology: Volume 2: Reality*. Wm. B. Eerdmans Publishing Co.
- McKirnan, D. J., & Hamayan, E. V. (1980). Language norms and perceptions of ethnolinguistic group diversity. In H. Giles, W. P. Robinson, & P. M. Smith (Eds.), *Language: Social psychological perspectives*. (pp. 161-69). Oxford: Pergamon Press.
- Meltzer, B. (2003). Lying: Deception in human affairs. *International Journal of Sociology and Social Policy*, vol. 23, no. 6-7, 61-79.
- Midgley, M. (1991). The origin of ethics. In P. Singer (Ed.), *A companion to ethics*. (pp. 3-13). Oxford, UK: Blackwell Publishers.
- Midgley, M. (2001). Why memes?. In H. Rose, & S. Rose (Eds.), *Alas poor Darwin: Arguments against evolutionary psychology*. London, UK: Random House.
- Millikan, R. (1984). *Language, thought, and other biological categories: New foundations for realism*. Cambridge, MA: MIT Press.
- Moll, J., de Oliveira-Souza, R., Eslinger, P., Bramati, I., Mourao-Miranda, J., Andreiuolo, P., et al. (2002). The neural correlates of moral sensitivity; A

- functional magnetic resonance imaging investigation of basic and moral emotions. *The Journal of Neuroscience*, 22:7, 2730-36.
- Monroe, K. (2002). Explicating Altruism. In S. G. Post, L. G. Underwood, J. P. Schloss, & W. B. Hurlbut (Eds.), *Altruism and altruistic love: Science, philosophy, and religion in dialog*. London, UK: Oxford University Press.
- Morain, G. G. (1987). Kinesics and cross-cultural understanding. In L. F. Luce, & E. C. Smith (Eds.), *Towards Internationalism*. Cambridge, MA: Newbury House Publishers.
- Motluk, A. (2002). You are what you speak. *New Scientist*, 176:2371, 34.
- Nelson, P. (1996). The Role of Theology in Current Evolutionary Reasoning. *Biology and Philosophy*, 11, 493-517.
- Newman, M., Pennebaker, J., Berry, D., & Richards, J. (2003). Lying Words: Predicting Deception from Linguistic Styles. *Personality and Social Psychology Bulletin*, vol. 29, no. 5, 665-75.
- Newport, E. (1997). Maturation constraints on language learning. *Cognitive Science*, 14, 11-28.
- Niemeier, S., & Dirven, R. (Eds.). (2000). *Evidence for Linguistic Relativity: Amsterdam studies in the theory and history of linguistic science. Series IV, Current issues in linguistic theory*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Norman, R. (1997). Making sense of moral realism. *Philosophical investigations*, 20(2), 117-35.
- Oates, D. (1988). Social Darwinism and Natural Theodicy. *Zygon*, 23:4, 439-54.
- Oliner, S. P. (2002). Extraordinary acts of ordinary people: Faces of heroism and altruism. In S. G. Post, L. G. Underwood, J. P. Schloss, & W. B. Hurlbut (Eds.), *Altruism and altruistic love*. London, UK: Oxford University Press.
- Onishi, M. (1997). The grammar of mental predicates in Japanese. *Language Sciences*, 19(3), 219-33.
- Orwell, G. (1950). 1984. Harmondsworth, Middlesex, UK: Penguin Books.
- Osgood, C., & Tannenbaum, P. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.
- Pinker, S. & Bloom, P. (1990). Natural language and natural selection. *Behavior and Brain Sciences*, 13, 707-84.
- Pinker, S. (1994). *The language instinct*. London, UK: Penguin Books.

- Pinker, S. (2002). *The blank slate: The modern denial of human nature*. London: Penguin Books.
- Plotkin, H. (2000). People Do More Than Imitate. *Scientific American*, 283, No 4, 72.
- Polkinghorne, J. (2003). *Belief in God in an Age of Science (Yale Nota Bene)*. Yale University Press.
- Post, S. G., Underwood, L. G., Schloss, J. P., & Hurlbut, W. B. (Eds.). (2002). *Altruistic and altruistic love*. London, UK: Oxford University Press.
- Povinelli, D. & Vonk, J. (2004). We don't need a microscope to explore the chimpanzee's mind. *Mind and Language*, Vol. 19 No. 1, 1-28.
- Povinelli, D., Bering, J., & Giambrone, S. (2000). Toward a science of other minds: Escaping the argument by analogy. *Cognitive Science*, Vol. 24(3), 509-41.
- Power, T. G. (2000). *Play and exploration in children and animals*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Premack, D. (2004). Is language the key to human intelligence? *Science*, 303, 318-20.
- Preston, C. (2002). Animality and morality: Human reason as an animal activity. *Environmental Values*, 11, 427-42.
- Preston, S. & de Waal, F. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25:1, 1-72.
- Pütz, M., & Verspoor, M. (Eds.). (2000). *Explorations in Linguistic Relativity. Amsterdam studies in the theory and history of linguistic science. Series IV, Current issues in linguistic theory*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Renshaw, D. C. (1993). Lies and medicine: reflections on the etiology, pathology, and diagnosis of chronic lying. *Clin Ther*, 15(2), 465-73; discussion 432.
- Richards, J. W., Gilder, G. F., Kurzweil, R., Ray, T., Searle, J., Dembski, W., et al. (2002). *Are We Spiritual Machines?: Ray Kurzweil vs. the Critics of Strong A.I.* Discovery Institute.
- Richerson, J., & Boyd, P. (1997). The evolution of human ultra-sociality. In I. Eibl-Eibesfeldt, & F. Salter (Eds.), *Ideology, warfare, and indoctrinability*.
- Rose, H., & Rose, S. (Eds.). (2001). *Alas poor Darwin: Arguments against evolutionary psychology*. London, UK: Random House.
- Rose, S. (1997). *Lifelines: Biology, freedom, determinism*. London, UK: Penguin Books.

- Ross, D., Thompson, R., & Yenkosky, J. (1997). Laterization of affective prosody in brain and the callosal integration of hemispheric language functions. *Brain and Language*, 56, 27-54.
- Ruse, M. (1991). The significance of evolution. In P. Singer (Ed.), *A companion to ethics*. (pp. 500-10). Oxford, UK: Blackwell Publishers.
- Ruse, M. (2002). A Darwinian naturalist's perspective on altruism. In S. G. Post, L. G. Underwood, J. P. Schloss, & W. B. Hurlbut (Eds.), *Altruistic and altruistic love*. (pp. 168-81). London, UK: Oxford University Press.
- Savage-Rumbaugh, S., Shanker, S. G., & Taylor, T. J. (2001). *Apes, Language, and the Human Mind*. Oxford University Press.
- Schloss, J. (1998). Evolutionary accounts of altruism & the problem of goodness by design. In W. Dembski (Ed.), *Mere creation science, faith and intelligent design*. (pp. 236-61). Downers Grove, Ill: InterVarsity.
- Schloss, J. (2002). Emerging accounts of altruism. In S. G. Post, L. G. Underwood, J. P. Schloss, & W. B. Hurlbut (Eds.), *Altruism and Altruistic Love*. London, UK: Oxford University Press.
- Schore, A. (1994). *Affect regulation and the origin of the self: The neurobiology of emotional development*. Hove, UK: Lawrence Erlbaum Associates.
- Schore, A. (1998). The Experience-dependent Maturation of an Evaluative System in the Cortex. In K. H. Pribaum (Ed.), *Brain and Values: Is a Biological Science of Values Possible*. (pp. 337-58). Mahwah, NJ: Lawrence Erlbaum Associates.
- Singer, P. (1981). *The expanding circle: ethics and sociobiology*. New York.: Farrar, Straus & Giroux.
- Singer, P. (1997). The drowning child and the expanding circle. *New Internationalist*. Retrieved February 13, 2006, <http://www.newint.org/issue289/drowning.htm>
- Smith, K. (2003). The Transmission of Language: models of biological and cultural evolution. Unpublished PhD, University of Edinburgh.
- Snow, C. (1990). Language and the beginnings of moral understanding. In J. Kagan, & S. Lamb (Eds.), *The emergence of morality in young children*. Chicago: University of Chicago Press.
- Sober, E. & Wilson, D. (1994). A Critical Review of the Philosophical Work on the Units of Selection Problem. *Philosophy of Science*, 61, 534-55.
- Sober, E. (2005). Sex Ratio Theory, Ancient and Modern — The Debate Among Arbuthnot, Bernoulli, and
- Sober, E., & Wilson, D. (1998). *Unto Others: The Evolution and Psychology of*

Unselfish Behavior. Harvard University Press.

- Sober, E., & Wilson, D. (2000). Summary of *Unto Others: The Evolution and Psychology of Unselfish Behavior*. In L. Katz (Ed.), *Evolutionary Origins of Morality: Cross-Disciplinary Perspectives*. UK: Thorverton Imprint Academic.
- Sommers, T. & Rosenberg, A. (2003). Darwin's nihilistic idea: evolution and the meaningless of life. *Biology and Philosophy*, 18, 653-68.
- Sorokin, P. A., & Post, S. G. (2002). *The Ways and Power of Love: Types, Factors, and Techniques of Moral Transformation*. Templeton Foundation Pr.
- Sperber, D. (2000). An objection to the memetic approach to culture. In R. Auger (Ed.), *Darwinizing culture: The status of memetics as science*. London: Oxford University Press.
- Spinka, M., Newberry, R., & Bekoff, M. (2001). Mammalian play: training for the unexpected. *Quarterly Review of Biology*, 76:2, 141-68.
- Spradley, J. P. (1979). *The ethnographic interview*. New York: Hold, Rinehart and Winston.
- Staub, E. (1992). *The Roots of Evil : The Psychological and Cultural Origins of Genocide*. Cambridge University Press.
- Stearns, S. C., & Hoekstra, R. F. (2000). *Evolution : An Introduction*. Oxford University Press, USA.
- Stiff, J. B., & Miller, G. R. (1993). *Deceptive Communication*. Sage Publications.
- Strungaru, C. A., & Schiefenhövel, W. (2002). The Word to the Wise: Proverbs and the Patterns of the Mind. *Science and Spirit*. Retrieved February 13, 2006, http://www.science-spirit.org/articles/articledetail.cfm?article_id=131
- Tappan, M. (1997). Language, culture, and moral development: A Vygotskian perspective. *Developmental Review*, 17, 78-100.
- Tattersall, I. (1998). *Becoming human: Evolution and human uniqueness*. London, UK: Oxford University Press.
- Terashima, S. (2003). Tension and tense: Japan and Japanese. Retrieved February 13, 2006, <http://www.u-ryukyu.ac.jp/old/nakayukui/old/105/6-2.html>
- The Scotsman. (2004). Congress Calls Rumsfeld to Account as Bush Says Sorry for Abuse. Retrieved May 7, 2004, www.scotsman.com
- Thierry, B. (2000). Building Elements of Morality Are Not Elements of Morality. In L. Katz (Ed.), *Evolutionary Origins of Morality: Cross-Disciplinary Perspectives*. UK: Thorverton; Imprint Academic.

- Trivers, R. (1971). The Evolution of Reciprocal Altruism. *Quarterly Review of Biology*, 46, 35-57.
- Trivers, R. (1985). *Social evolution*. Menlo Park, California: The Benjamin/Cummings Publishing Company.
- Underwood, L. (2002). The human experience of compassionate love: Conceptual mapping and data from selected studies. In S. G. Post, L. G. Underwood, J. P. Schloss, & W. B. Hurlbut (Eds.), *Altruism and altruistic love*. London, UK: Oxford University Press.
- Van Lancker, D. (1997). Rags to riches: Our increasing appreciation of cognitive and communicative abilities of the human right cerebral hemisphere. *Brain and Language*, 57, 1-11.
- Vine, I. (2000). Selfish, Altruistic, or Groupish? Natural Selection and Human Moralities. In L. Katz (Ed.), *Evolutionary Origins of Morality: Cross-Disciplinary Perspectives*. UK: Thorverton; Imprint Academic.
- Voland, E. & Voland, R. (1995). *Journal of Social and Evolutionary Systems*, 18(4), 397-412.
- Vygotsky, L. (1981). The instrumental method in psychology. In J. Wertsch (Ed.), *The concept of activity in Soviet psychology*.
- Wallace, W., T. (1994). Memory for Music: Effect of Melody on Recall of Text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(6), 1471-85.
- Waller, M. J. C. (1995). Darwinism and the enemy within. *Journal of Social and Evolutionary Systems*, 18(3), 217-29.
- Weikart, R. (2004). *From Darwin to Hitler : Evolutionary Ethics, Eugenics, and Racism in Germany*. Palgrave Macmillan.
- Weisberg, J. (Ed.). (2001). *George W. Bushisms*. New York, NY: Fireside.
- Wierzbicka, A. (1992). Defining emotion concepts. *Cognitive Science*, 16(4), 539-81.
- Wierzbicka, A. (1994). Semantics and epistemology: The meaning of 'evidentials' in a cross-linguistic perspective. *Language Sciences*, 16(1), 81-137.
- Wierzbicka, A. (1996). *Semantics: Primes and universals*. Oxford, UK: Oxford University Press.
- Wikipedia. (2003). Wikipedia: The Free Encyclopedia: The Golden Rule. Retrieved February 13, 2006, http://www.wikipedia.org/wiki/Golden_rule
- Williams, G. (1996). Adaptation and Natural Selection: A Critique of Some Current

Evolutionary Thought. Princeton University Press.

Wilson, D. (2002). *Darwin's cathedral: Evolution, religion, and the nature of society*. Chicago, IL: University of Chicago Press.

Wilson, D. S., Wilczynski, C., Wells, A., & Weiser, L. (2000). Gossip and other aspects of language as group-level adaptations. In C. Heyes, & L. Huber (Eds.), *The evolution of cognition*. (pp. 347-65). Cambridge, Massachusetts: MIT Press.

Wilson, D. (1999). Flying over uncharted territory. *Science*, 285 (5425), 206.

World Vision. (2003). 2003 Annual Review. *World Vision International*.

Wortham, S. & Locher, M. (1999). Embedded metapragmatics and lying politicians. *Language & Communication*, 19(2), 109-25.

Wray, A. (2002). Dual processing in protolanguage; Performance without competence. In A. Wray (Ed.), *The transition to language*. London: Oxford University Press.

Wray, A. (Ed.). (2002). *The transition to language*. London, UK: Oxford University Press.

Wyschogrod, E. (2002). Pythagorean bodies and the body of altruism. In S. G. Post, L. G. Underwood, J. P. Schloss, & W. B. Hurlbut (Eds.), *Altruism and altruistic love*. London, UK: Oxford University Press.

Yalch, R., F. (1991). Memory in a Jingle Jungle: Music as a Mnemonic Device in Communicating Advertising Slogans. *Journal of Applied Psychology*, 76(2), 268-75.

Yanagisawa, M. (2002). Personal communication.

Zahavi, A., & Zahavi, A. (1999). *The Handicap Principle: A Missing Piece of Darwin's Puzzle*. Oxford University Press.